Subdirect products of free semigroups and monoids

Ashley Clayton

University of
St Andrews

This thesis is submitted in partial fulfilment for the degree of
Doctor of Philosophy (PhD)
at the University of St Andrews

April 2020

Abstract

Subdirect products are special types of subalgebras of direct products. The purpose of this thesis is to initiate a study of combinatorial properties of subdirect products and fiber products of semigroups and monoids, motivated by the previous work on free groups, and some recent advances in general algebra.

In Chapter 1, we outline the necessary preliminary definitions and results, including elements of algebraic semigroup theory, formal language theory, automata theory and universal algebra.

In Chapter 2, we consider the number of subsemigroups and subdirect products of $\mathbb{N} \times \mathbb{N}$ up to isomorphism. We obtain uncountably many such objects, and characterise the finite semigroups S for which $\mathbb{N} \times S$ has uncountably many subsemigroups and subdirect products up to isomorphism.

In Chapter 3, we consider particular finite generating sets for subdirect products of free semigroups introduced as "sets of letter pairs". We classify and count these sets which generate subdirect and fiber products, and discuss their abundance.

In Chapter 4, we consider finite generation and presentation for fiber products of free semigroups and monoids over finite fibers. We give a characterisation for finite generation of the fiber product of two free monoids over a finite fiber, and show that this implies finite presentation. We show that the fiber product of two free semigroups over a finite fiber is never finitely generated, and obtain necessary conditions on an infinite fiber for finite generation.

In Chapter 5, we consider the problem of finite generation for fiber products of free semigroups and monoids over a free fiber. We construct two-tape automata which we use to determine the language of indecomposable elements of the fiber product, which algorithmically decides when they are finitely generated.

Finally in Chapter 6, we summarise our findings, providing some further questions based on the results of the thesis.

Declarations

Candidate's declaration

I, Ashley Clayton, do hereby certify that this thesis, submitted for the degree of PhD, which is approximately 40 000 words in length, has been written by me, and that it is the record of work carried out by me, or principally by myself in collaboration with others as acknowledged, and that it has not been submitted in any previous application for any degree.

I was admitted as a research student at the University of St Andrews in September 2016.

I received funding from an organisation or institution and have acknowledged the funder(s) in the full text of my thesis.

Date 30/07/20 Signature of candidate

Supervisor's declaration

I hereby certify that the candidate has fulfilled the conditions of the Resolution and Regulations appropriate for the degree of PhD in the University of St Andrews and that the candidate is qualified to submit this thesis in application for that degree.

Date 30/07/20 Signature of supervisor

Permission for publication

In submitting this thesis to the University of St Andrews we understand that we are giving permission for it to be made available for use in accordance with the regulations of the University Library for the time being in force, subject to any copyright vested in the work not being affected thereby. We also understand, unless exempt by an award of an embargo as requested below, that the title and the abstract will be published, and that a copy of the work may be made and supplied to any bona fide library or research worker, that this thesis will be electronically accessible for personal or research use and that the library has the right to migrate this thesis into new electronic forms as required to ensure continued access to the thesis.

I, Ashley Clayton, confirm that my thesis does not contain any third-party material that requires copyright clearance.

The following is an agreed request by candidate and supervisor regarding the publication of this thesis:

Printed copy

No embargo on print copy.

Electronic copy

No embargo on electronic copy.

Date 30/07/20 Signature of Candidate

Date 30/07/20 Signature of Supervisor

Underpinning Research Data or Digital Outputs

Candidate's declaration

I, Ashley Clayton, hereby certify that no requirements to deposit original research data or digital outputs apply to this thesis and that, where appropriate, secondary data used have been referenced in the full text of my thesis.

Date 30/07/20 Signature of candidate

Contents

Abstract **i**

1 Introduction and preliminaries **1**

1.1 Conventions and structure 7
1.2 Semigroups, monoids and homomorphisms 8
1.3 Relations, congruences, quotients of semigroups & monoids . . 18
1.4 Green's relations on semigroups and monoids 22
1.5 Formal language theory, automata, free semigroups & monoids 27
1.6 Generating and presenting semigroups and monoids 30
1.7 Direct products, subdirect products and fiber products 41

2 Subdirect products involving the free monogenic semigroup **50**

2.1 Subsemigroups of $\mathbb{N} \times \mathbb{N}$ 52
2.2 Subsemigroups of direct products of $\mathbb{N}$ with a finite semigroup 64
2.3 Subdirect powers of $\mathbb{N}$ 70
2.4 Subdirect products of $\mathbb{N}$ with a finite semigroup 72

3 Counting finitely generated subdirect products and fiber products of free semigroups **78**

3.1 Sets of letter pairs generating subdirect products of free semigroups 79
3.2 Sets of letter pairs generating fiber products of free semigroups 84
3.3 Proportion of sets of letter pairs generating subdirect products and fiber products 91

4 Finitary properties for fiber products of free semigroups and monoids 96

4.1 Finite generation for fiber products of free monoids over finite fibers 98

4.2 Finite presentation for fiber products of free monoids over finite fibers 107

4.3 Finite generation for fiber products of free semigroups over infinite fibers 118

5 Deciding finite generation for fiber products of free semigroups and monoids 131

5.1 Equivalent conditions to finite generation for fiber products of free semigroups and monoids over free fibers 133

5.2 Two-tape automata construction for fiber products of free monoids over free monoid fibers 136

5.3 The language recognised by the two-tape automaton associated with $\Pi(\varphi, \psi)$ 144

5.4 Decidability of finite generation for fiber products of free semigroups and monoids over free fibers 158

6 Concluding remarks and further questions 163

Bibliography 167

Acknowledgements 170

Chapter 1

Introduction and preliminaries

In its operational form, the algebraic theory of semigroups is in part the developing abstraction of the algebraic theory of groups. The latter theory has of course been well studied and researched as early as the late 18th century. Whereas elements of group theory have formalised the notions of symmetries for mathematical objects, semigroup theory attempts to formalise notions of *partial symmetries* and *transformations* of those objects. A particularly notable example of this is the *Wagner-Preston theorem* ([16, Theorem 5.1.7]), which states that any *inverse* semigroup (called "generalised" groups, by Wagner [29]) is embeddable in a *symmetric inverse semigroup* (a "generalised" symmetric group). The symmetric inverse semigroup is a structure of partial symmetries (see [16, Theorem 5.1.5]), and so the Wagner-Preston theorem gives the partial symmetry analogue to Cayley's theorem for groups.

Though the origins of the theory of semigroups can be traced back to the early twentieth century in the papers of Dickson [11] and de Séguier [10], it has been argued that its research beginnings can be attributed with Sushkevich [28] in the late 1920s, and the celebrated paper of Rees [25] in 1940. This makes it a relatively recent and modern study in the history of Algebra, and Mathematics in general. In this author's opinion, the theory benefits from its youth with a metaphorical continent of roads untravelled. Though it

comprises many interesting open problems in its own right, it further provides many theoretical techniques and insights into other algebraic studies, such as group theory. Moreover, it has seen practical application via the related study of theoretical computer science, which is another relatively recent field. One closely linked brethren theory to that of semigroup theory is the study of *finite automata and formal languages*. This study encompasses the theoretical formalisation of algorithmic decision making machines, such as finite state automata, transducers, and Turing machines. There is a plethora of practical computing application in this field, though its formulisation lends itself to the study of pure semigroup theory particularly well. Strings of characters can be viewed as abstract semigroup products of elements referred to as *letters*, and vice versa. These products can be viewed as *words* over the *alphabet* of the possible characters for a string. The set of all such words form *languages*, which can be *recognised* by the machinery of automata, which can decide if a given word is in a language or not. As such, problems of deciding properties of semigroup elements have been formulated as *decision problems* in the theory of automata.

Though this work comprises many elements of the pure and applied studies outlined above, it is primarily motivated from the perspective of pure semigroup theory. Its nature then will be to investigate, develop, and advance theoretical concepts of semigroup theory from the metaphorically aforementioned roads less travelled. In this work, that road will be the theory of subdirect products of semigroups. Heuristically, in the more general theory of universal algebra, subdirect products are special types of subalgebras of direct products of algebras. Each direct factor can in some way be embedded into the subdirect product, making the factors of the subdirect product "full" in some sense. As substructures of direct products, subdirect products in a way generalise the algebraic construction of direct products. However, we will consider being a subdirect product more of a property of an algebra, rather than a construction.

In the latter half of the 20th century, subdirect products have been notably studied in the algebraic theory of groups. Some studies have taken a com-

binatorial approach, such as in the work of McKenzie [20]. Particularly in this work, the number of subdirect products (up to isomorphism) of a countable power of a finite group G is found to be countable precisely when G is abelian. In a similar question of countability, Bridson & Miller find uncountably many subdirect products of two free groups up to isomorphism [3, Corollary B]. In addition to this, they (as well as other authors such as Baumslag & Roseblade [2] and Mikhaĭlova [21]) focus on classical questions relating to *finitary properties* (those which hold for all finite groups) of subdirect products. Namely, questions such as "when are subdirect products of groups *finitely generated*, *finitely presented*, when do they have *decidable membership problem*?" amongst others are asked and answered.

It has been a common trend in the relevant literature to provide examples of subdirect products of groups with particular combinations of finitary properties. Notably, Mikhaĭlova [21] uses subdirect products of two free groups to provide examples of finitely generated groups with undecidable membership problem. Further, Bridson & Miller give subdirect products which are not finitely generated [3, Example 3], and Grunewald gives subdirect products which are finitely generated but not finitely presented [13, Proposition B]. We remark that all of this work utilises subdirect products of free groups in particular.

A way of constructing such examples of subdirect products is via *fiber products*, also known as *pullbacks* in the theory of categories. Fiber products are constructed from two (or more) group epimorphisms onto some common image. Elements of the domain groups are paired together when their images coincide (or placed in a n-tuple when there are n epimorphisms), and the resulting structure (using the multiplication from the direct product) gives a subdirect product. For groups, fiber products and subdirect products are one and the same; every subdirect product of groups is constructible in this way. This is a consequence of a lemma of Goursat [14, Theorem 5.5.1], but more generally can be derived by a lemma of Flesicher (given in [4, Lemma 10.1]). Fleischer's lemma states that a subdirect product is a fiber product if and only if the kernels of the projection maps onto each coordinate commute

as algebraic congruences. As a consequence, subdirect products and fiber products coincide in given *varieties* of algebras (which are special classes of algebras) precisely when the congruences on those varieties commute. This is the case for the variety of groups, but not for those of semigroups, monoids, and other algebraic structures.

Recently, Mayr & Ruškuc [19] have explored subdirect products and fiber products from the viewpoint of universal algebra, with a particular emphasis on finitary properties. Within this work, they give an example of a subdirect product of monoids which is not finitely generated [19, Example 7.1], and an example of a fiber product of monoids which is finitely presented but with non-finitely presented factors [19, Example 7.3]. These examples in particular utilise free monoids, in comparison to the work of Bridson & Miller. Further works involving subdirect products of semigroups have been provided by Chrislock & Tamura [5] (with focus on products involving rectangular bands), Nambooripad & Veeramony [23] (for regular semigroups), Mitsch [22] (for E-inversive semigroups), amongst others. These works all have specialised their focus. A more general treatment of the theory of subdirect products of semigroups, leading on from the free group works of the aforementioned authors, has yet to be undertaken in the literature.

The purpose of this thesis is to initiate a study of subdirect products and fiber products of free semigroups and monoids, motivated by the work for subdirect products of free groups aforementioned. Our primary aims will be to place particular emphasis on finitary properties such as finite generation and presentation, though we will begin by exploring combinatorial questions of isomorphism, motivated by Bridson & Miller [3]. We will argue the narrative that finitely generated subdirect products of free semigroups and monoids are easily found, yet fiber products of free semigroups and monoids are not easily finitely generated, if at all. This will highlight the difference in nature between fiber products and subdirect products for semigroups.

In the rest of Chapter 1, we introduce the planned structure of this thesis, and outline the necessary preliminary definitions and results needed. We will

begin by covering the introductory concepts of semigroup theory utilised in this work. Such concepts include semigroup isomorphisms, quotients, congruences, Green's relations, finite generation for semigroups and semigroup presentations. We will also introduce the theory of formal language and automata, which will play an intrinsic role in our discussion of deciding finite generation later in the thesis. Finally, alongside direct products, we will formalise the definitions of subdirect products and fiber products for semigroups. This will be accompanied by a spotlight on some of the examples from group theory outlined above, and the more general concepts and results from subdirect products in universal algebra, such as Flesicher's lemma.

In Chapter 2, we begin our study of subdirect products of free semigroups with a combinatorial review of subdirect products involving the free monogenic semigroup, $\mathbb{N}$. Motivated by Bridson & Miller [3], our main aim will be to show that the direct product $\mathbb{N} \times \mathbb{N}$ contains uncountably many subsemigroups up to isomorphism, of which uncountably many are subdirect products. We will accomplish this by constructing families of finitely generated subdirect products using 3-separating sets, which we introduce and define. As a corollary to this, we more generally highlight that the direct product of two semigroups, each with elements of infinite order, has uncountably many subsemigroups up to isomorphism. We give analogous results for the finite direct power $\mathbb{N}^k$, and further consider the number of non-isomorphic subsemigroups and subdirect products of $\mathbb{N} \times S$, where S is a finite semigroup. In particular, we show that $\mathbb{N} \times S$ has countably many non-isomorphic semigroups if and only if S is completely regular, and has countably many non-isomorphic subdirect products if and only if every element of S has a relative left or right identity.

In Chapter 3, we begin our considerations of subdirect products and fiber products of free semigroups of rank higher than one, transitioning into questions of finite generation. We lead from the combinatorial motivations of Chapter 2, outlining our aim to count the number of particular finite subsets (introduced as *sets of letter pairs*) generating subdirect products and fiber products of two free semigroups. In particular, given finite alphabets A and

B, we determine the number of subsets of $A \times B$ generating subdirect products of $A^+ \times B^+$, and fiber products of $A^+ \times B^+$ as separate results. We give some analytic assessment of these results, in particular arguing that the ratio of subsets of $A \times A$ generating subdirect products of $A^+ \times A^+$ approaches 1 as $|A|$ grows. We contrast this by arguing that the ratio of subsets of $A \times A$ generating fiber products of $A^+ \times A^+$ approaches 0 as $|A|$ grows. We hence comment on the difference in nature between generating fiber products of free semigroups and subdirect products of free semigroups to conclude the chapter.

In Chapter 4, we transition to questions of finite generation and presentation for fiber products of semigroups and monoids. In paticular, we aim to characterise finite generation for the fiber products of two free monoids over a finite fiber, and separately for the fiber products of two free semigroups over a finite fiber. In the latter case, we show that no such fiber products are finitely generated, but completely characterise the former case. We do this by showing that the given epimorphisms of the fiber product map every element of the free monoid alphabet to the same element of the fiber, which must be a cyclic group. In particular for this case, we show that finite generation implies finite presentation for fiber products of free monoids over finite fibers, by directly finding presentations for them. Returning to the case for fiber products of free semigroups, we consider necessary conditions on the fiber for them to be finitely generated. Namely, we show the fibers must be infinite, finitely generated, $\mathcal{J}$-trivial, idempotent free semigroups. We conclude the chapter by showing that these conditions, though restrictive, are not sufficient for finite generation, using the examples of free commutative semigroups.

In Chapter 5, we lead from the findings of Chapter 4 by initiating a technical study for deciding finite generation for fiber products of free semigroups and monoids. We begin by noting that such fiber products are finitely generated precisely when they contain finitely many indecomposable elements. We then introduce the machinery of *two-tape automata* from formal language theory, in order to attempt to recognise the language of these indecomposable

elements. Specifically, for a fiber product of two free monoids over a free fiber, we construct an associated two-tape automaton which recognises a language in bijection with the indecomposable elements of the fiber product. We then give sufficient and necessary conditions for finite generation of the fiber product, in terms of finding cycles within the associated automaton. We show as a corollary that the problem of finite generation for fiber products of free monoids over free fibers is decidable. Finally, we give the analogous results to those previously described in the chapter, for fiber products of free semigroups over free fibers.

We conclude the thesis in Chapter 6, with a brief discussion summarising our findings. We give particular attention to some further questions relating to material in this thesis, and argue that they provide many fruitful points of continuation for the study of subdirect products of semigroups and monoids.

1.1 Conventions and structure

This thesis has been divided into chapters and sections. Theorems, lemmas, corollaries etc have been numbered as **x.y.z**, where **x** is the chapter of the result, **y** is the section number relative to that chapter, and **z** is the order of the result within the section. The ends of proofs, examples, and definitions/notation will be indicated by the symbols $\square$, $\triangle$ and $\blacksquare$ respectively.

The reader is assumed to have a basic knowledge of group theory at the undergraduate level. Throughout, we will by convention refer to the natural numbers $\mathbb{N}$ as the set $\{1, 2, 3, \dots\}$ **not** containing zero. The set $\{0, 1, 2, 3, \dots\}$ will be denoted by $\mathbb{N}^0$. Whereas other authors may use the term *countable* to strictly mean *countably infinite* (that is, the cardinality of a set in bijection with $\mathbb{N}$), we will use the term countable to mean either countably infinite, or finite.

For the purposes of composition the usual convention in semigroup theory is

to write maps to the right of the argument. However as composition of maps does not appear as a main feature of this thesis, maps will be written on the left of the argument by convention (e.g.. $f(x)$). Wherever composition does feature however, we will use the convention that

$$(f \circ g)(x) = g(f(x)),$$

so maps are indeed composed from left to right as read.

1.2 Semigroups, monoids and homomorphisms

In this section, we introduce the necessary concepts and definitions from the algebraic theory of semigroups used in this work. However, we begin by recalling a couple of brief definitions from the theory of universal algebra, namely those of an *algebra* and a *variety*. A comprehensive study of these concepts may be found in [4].

Definition 1.2.1. An *algebra* is a set A together with a collection of n-ary operations on A (maps taking n elements of A as inputs, to a single output of A).

The *type* of an algebra is an ordered sequence of the natural numbers n comprising the n-ary operations on A. This ordering is usually from largest to smallest. ■

Definition 1.2.2. A *variety* of algebras is a class of algebras of the same type, which is closed under taking subalgebras, direct products, and homomorphic images. ■

Example 1.2.3. A group G comes equipped with the binary operation of multiplication, the unary operation of inversion, and the nullary operation which quantifies the existence of the identity. Hence a group can be described as an algebra of type $(2, 1, 0)$.

As any subgroup of a group is a group, the direct product of two groups is a group, and the group homomorphic image of any group is again a group,

then the class of groups forms a variety. △

We now begin with the basic definitions and results required from the theory of semigroups. We kindly refer the reader to sources such as [16], [8], and [15] for further comprehensive studies.

Definition 1.2.4. A *semigroup* is a pair $(S, \cdot)$, where S is a non-empty set, and $\cdot : S \times S \to S$ is a binary operation that is also *associative*, meaning for all $s, t, u \in S$,

$$s \cdot (t \cdot u) = (s \cdot t) \cdot u.$$ ■

Where the context of the operation is clearly understood, we will speak of the semigroup S only by referring to its set. We will also omit the symbol $\cdot$, and instead write st for $s \cdot t$, which we will call the *product* of s and t. We will further talk about the operation as the *multiplication* of S. The multiplicative notation allows us to consider *powers* of elements s^n for $n \in \mathbb{N}$, without amiguity of meaning. When S is finite, we can use a *multiplication table* to descibe the multiplication of S, similarly to finite groups.

Examples 1.2.5. • The set $\{0, 1\}$ together with the usual multiplication of real numbers forms a *finite* semigroup;

• The natural numbers $\mathbb{N}$ together with addition form a semigroup;

• The natural numbers $\mathbb{N}$ together with subtraction does *not* form a semigroup, as $1 - (2 - 3) = 2$, but $(1 - 2) - 3 = -4$. △

As all groups have an associative binary operation, then every group is of course a semigroup. Every group also has an identity element, and hence is also a *monoid*, which we now define.

Definition 1.2.6. A *monoid* is a semigroup M containing an *identity* $1 \in M$, meaning for all $m \in M$,

$$1m = m1 = m.$$ ■

It is a short exercise to see that identities in monoids are unique, hence we

will always refer to *the* identity of M.

Examples 1.2.7. • The finite semigroup $(\{0,1\},\times)$ is also a finite monoid, with 1 being the identity;

• $(\mathbb{N},+)$ does *not* form a monoid, but $\mathbb{N}^0$ with addition *does*, taking 0 to be the identity;

• Every group is a monoid (and so every group is also a semigroup). △

The identity 1 of a monoid M has the property that $1^2 = 1$. The identity need not be the only element with this property however, as seen in the following definition and examples.

Definition 1.2.8. An *idempotent* of a semigroup S is an element $e \in S$ such that $e^2 = e$. The set of all idempotents of a semigroup S will be denoted by $E(S)$. A semigroup is said to be *idempotent-free* if $E(S) = \emptyset$. ■

Idempotents hence also satisfy $e^{n+1} = e^n$ for all $n \in \mathbb{N}$.

Examples 1.2.9. • $(\mathbb{N},+)$ has no idempotents, as $n + n = 2n$, which is strictly larger than n;

• $(\mathbb{N}^0,+)$ has one idempotent 0, as $0 + 0 = 0$;

• Both elements of $(\{0,1\},\times)$ are idempotents, as $0^2 = 0$ and $1^2 = 1$. △

We next see that idempotents always exist in finite semigroups.

Lemma 1.2.10. *Let S be a semigroup, and let $s \in S$. If S is finite, then there is some $k \in \mathbb{N}$ such that s^k is idempotent. In particular, $E(S) \neq \emptyset$.*

Proof. Let $s \in S$, and consider the set of all powers of s, given by $T = \{s, s^2, s^3, \dots\}$. As S is finite, then T must also be finite, and hence there must exist some $m \in \mathbb{N}$ such that $T = \{s, s^2, s^3, \dots, s^m\}$. In particular, considering the element s^{m+1}, then there must be some $n \in \{1, 2, \dots, m\}$ such that $s^{m+1} = s^n$.

If $m = 1$, then $s^2 = s$, and s is idempotent and the result follows. Otherwise, as $m + 1 > n$, we can write $m + 1 = n + l$ for some $l \in \mathbb{N}$. As

$$s^n = s^{m+1} = s^{n+l}, \tag{1.1}$$

then by repeatedly multiplying (1.1) by s^l, it follows that $s^n = s^{n+jl}$ for all $j \in \mathbb{N}$. In particular then,

$$s^n = s^{n+nl} = s^{n(l+1)}. \tag{1.2}$$

If $l = 1$, then $s^n = s^{2n} = (s^n)^2$, and so s^n is idempotent. Otherwise, $l > 1$, and we can multiply both sides of (1.2) by $s^{n(l-1)}$, to get

$$\begin{aligned} & s^n s^{n(l-1)} = s^{n(l+1)} s^{n(l-1)} \\ \Rightarrow\ & s^{n+nl-n} = s^{nl+n+nl-n} \\ \Rightarrow\ & s^{nl} = s^{2nl} \\ \Rightarrow\ & s^{nl} = (s^{nl})^2. \end{aligned}$$

Hence s^{nl} is idempotent, and the result follows. □

As seen in the following result, a finite group can be realised as a finite monoid with a unique idempotent.

Lemma 1.2.11. *A finite monoid M is a group if and only if M has a unique idempotent.*

Proof. ($\Rightarrow$) If M is a group with an idempotent m, then $m^2 = m$, implying $m = 1$. Hence the identity is the unique idempotent of M.

($\Leftarrow$) Suppose M is not a group. Then there exists some $m \in M$ with no inverse. As M is finite, there are two distinct powers of m which are equal, i.e. $m^p = m^q$ for some $p, q \in \mathbb{N}$ with $p \neq q$. Assume without loss of generality that $p > q$ so that $p - q \in \mathbb{N}$. In particular, as

$$m^p = m^{p-q+q} = m^{p-q} m^q = m^{p-q} m^p = m^{p+(p-q)},$$

then p can be chosen large enough so we can further assume without loss of generality that $p > 2q$, so that $p - 2q \in \mathbb{N}$. Hence

$$(m^{p-q})^2 = m^{2p-2q} = m^{p-2q}m^p = m^{p-2q}m^q = m^{p-q},$$

and thus m^{p-q} is an idempotent. Further, m^{p-q} cannot be the identity of M, as m is assumed to have no inverse. Hence M has more than one idempotent. □

In the next definition, we note that the property of *commutativity* in abelian groups is formalised for semigroups as well.

Definition 1.2.12. A semigroup S is said to be *commutative* if $st = ts$ for all $s, t \in S$. ■

Of course, when S is a group, we will still say that S is "abelian" rather than commutative.

Examples 1.2.13. • $(\mathbb{N}, +)$ is a commutative semigroup;

• $(\{0, 1\}, \times)$ is a commutative semigroup;

• Define $\star$ on $\mathbb{N}$ by $m \star n = m$ for all $m, n \in \mathbb{N}$. Then

$$m \star (n \star o) = m \star n = m = m \star (n \star o),$$

and hence $(\mathbb{N}, \star)$ is a semigroup. However, it is not commutative, as $1 \star 2 = 1$, but $2 \star 1 = 2$. △

Given a group G, one can ask about the smaller *substructures* contained within it; often, we ask about its *subgroups*. In this thesis, we will place emphasis on studying the substructures of a semigroup, for the same reasons as one does for groups. Hence we give the following definitions of subsemigroups and submonoids.

Definition 1.2.14. A *subsemigroup* of a semigroup S is a subset $T \subseteq S$ such that T is closed under the multiplication of S. That is, for all $s, t \in T$,

$st \in T$. We will use the notation $T \leq S$ to indicate that T is a subsemigroup of S.

Similarly, a *submonoid* of a monoid M (with identity $1 \in M$) is a subsemigroup $N \subseteq M$ such that $1 \in N$. ■

Examples 1.2.15. • $\mathbb{N}$ is a subsemigroup of $\mathbb{N}^0$, as $\mathbb{N}$ is closed under addition;

• $\{0\}$ is a submonoid of $\mathbb{N}^0$, as $0 + 0 = 0$; △

A notable example of subsemigroups are *ideals* of semigroups, which we define after giving the following notation.

Notation 1.2.16. For a semigroup S, and subsets $X, Y \subseteq S$, define

$$XY = \{xy \in S : x \in X, y \in Y\}.$$

When X (or Y) is a singleton set $\{a\}$ for some $a \in S$, we will write aY (or Xa) instead of $\{a\}Y$ (or $X\{a\}$). ■

With this notation, for $X, Y, Z \subseteq S$ subsets of a semigroup S, we note that

$$\begin{aligned}(XY)Z &= \{tz \in S : t \in XY, z \in Z\} \\ &= \{(xy)z \in S : x \in X, y \in Y, z \in Z\} \\ &= \{x(yz) \in S : x \in X, y \in Y, z \in Z\} \\ &= \{xt \in S : x \in X, t \in YZ\} \\ &= X(YZ).\end{aligned}$$

Hence the power set $\mathcal{P}(S)$ becomes a semigroup with the above notation. We now define ideals of semigroups using this notation.

Definition 1.2.17. A *left ideal* of a semigroup S is a non-empty subset $I \subseteq S$ such that $SI \subseteq I$. Similarly, a *right ideal* is a non-empty subset $I \subseteq S$ such that $IS \subseteq I$. Finally, an *ideal* of a semigroup S is a non-empty subset $I \subseteq S$ which is both a left and right ideal of S. ■

Noting that $I^2 \subseteq SI$ and $I^2 \subseteq IS$, then we see that all ideals are also subsemigroups.

Examples 1.2.18. • Take the subset $I = \{x \in \mathbb{N} : x \geq 2020\}$ of $(\mathbb{N}, +)$. Let $m \in \mathbb{N}$, and $x \in I$. Then $x + m \geq 2020$, and hence $x + m \in I$. Thus as x and m were arbitrarily chosen, it follows that $I\mathbb{N} \subseteq I$, and hence I is a right ideal. Moreover as $x + m = m + x$, then I is also a left ideal, and thus an ideal of $\mathbb{N}$;

• Let $\star$ be the operation on $\mathbb{N}$ defined in Examples 1.2.13, and take $I \subseteq \mathbb{N}$ any non-empty subset, such that $I \neq \mathbb{N}$. Then for any $m \in \mathbb{N}$ and $x \in I$, it follows that

$$x \star m = x \in I,$$

and hence as m, x were arbitrary, then $I\mathbb{N} \subseteq I$ and so I is a right ideal of $\mathbb{N}$. As $I \neq \mathbb{N}$, take any $m \in \mathbb{N} \setminus I$. Then for all $x \in I$,

$$m \star x = m \notin I,$$

and hence $\mathbb{N}I \not\subseteq I$, so I is not a left ideal. Hence only left ideal of $(\mathbb{N}, \star)$ is $\mathbb{N}$ itself, and thus $\mathbb{N}$ is also the only two-sided ideal. △

We now focus on constructing particular *principal* ideals of a semigroup S from a given element $a \in S$. These ideals will become relevant when we define *Green's relations* later in Section 1.4.

Definition 1.2.19. Let S be a semigroup, and let $a \in S$. Define the following subsets of S:

$$Sa := \{sa : s \in S\},$$
$$aS := \{as : s \in S\},$$
$$SaS := \{sat : s, t \in S\}.$$

Note that these definitions respect the notation given in Notation 1.2.16.

Further, let 1 be a symbol not already belonging to S. Define the set

$$S^1 = \begin{cases} S & \text{if } S \text{ is a monoid,} \\ S \cup \{1\} & \text{otherwise.} \end{cases}$$

Then extending the multiplication of S to $S \cup \{1\}$ by defining

$$s1 = s,\ 1s = s, 11 = 1$$

for all $s \in S$, S^1 is a monoid, and is called the *monoid obtained from S by adjoining an identity if necessary.* It is an exercise left to the reader to see that

$$\begin{aligned} S^1a &= Sa \cup \{a\}, \\ aS^1 &= aS \cup \{a\}, \\ S^1aS^1 &= SaS \cup aS \cup Sa \cup \{a\}. \end{aligned}$$ ■

Definition 1.2.20. Let S be a semigroup, and let $a \in S$. The set S^1a is called the *principal left ideal generated by a.*

Similarly, the set aS^1 is called the *principal right ideal generated by a*, and the set S^1aS^1 is called the *principal two-sided ideal generated by a.* ■

Examples 1.2.21. • The subset $I = \{x \in \mathbb{N} : x \geq 2020\}$ of $\mathbb{N}$ from Examples 1.2.18 is an ideal of $S = \mathbb{N}$. We can consider the set S^1 by adjoining the symbol 0 to $\mathbb{N}$ (as $\mathbb{N}$ already contains a symbol 1). In this way, S^1 is exactly the set $\mathbb{N}^0$.

Letting $a = 2020$, then

$$aS^1 = \{a + m : m \in \mathbb{N}^0\} = \{2020 + m : m \in \mathbb{N}^0\}.$$

Every element of $x \in aS^1$ is such that $x \geq 2020$, and moreover any $x \geq 2020$ has some $m \in \mathbb{N}^0$ such that $x = 2020 + m$. Therefore $aS^1 = I$, and hence I is a principal right ideal of S, generated by 2020. It is also a principal left ideal (noting that addition is commutative), and is hence a principal two-sided

ideal as well.

The above argument also holds for any $a \in \mathbb{N}$, and hence the principal left, right, and two-sided ideals generated by a are given by

$$S^1a = aS^1 = S^1aS^1 = \{x \in \mathbb{N} : x \geq a\}.$$

• Let $(\mathbb{N}, \star)$ be the semigroup from Examples 1.2.18. We saw that any non-empty subset of $S = \mathbb{N}$ was a right ideal. As in the previous example, S^1 can be taken as the set $\mathbb{N}^0$. For all $a \in \mathbb{N}$, we have that

$$aS^1 = \{a \star m : m \in \mathbb{N}^0\},$$

but $a \star m = a$ for all $m \in \mathbb{N}^0$. Hence $aS^1 = \{a\}$, and so any singleton subset is a principal right ideal generated by its single element. We also saw that the only left and two-sided ideal of $(\mathbb{N}, \star)$ is $\mathbb{N}$ itself, and so every principal left ideal S^1a and every principal two-sided ideal S^1aS^1 must be equal to $\mathbb{N}$. △

We will return to these principal ideals in Section 1.4. We now move from discussing substructures of semigroups to maps that preserve structure. In group theory, recall that a group homomorphism is a mapping between two groups which respects the multiplication and inversion of the groups. This notion also generalises to *semigroup homomorphisms*, as seen in the next definition.

Definition 1.2.22. For two semigroups S and T, a *semigroup homomorphism* is a function $\varphi : S \to T$ respecting the multiplication of each semigroup. That is, for all $s, t \in S$,

$$\varphi(st) = \varphi(s)\varphi(t).$$

If S and T are two monoids with identities 1_S and 1_T respectively, then a *monoid homomorphism* is a function $\varphi : S \to T$ satisfying the above, with

the added condition that

$$\varphi(1_S) = 1_T.$$

A surjective homomorphism will be called an *epimorphism*, an injective homomorphism will be called a *monomorphism*, and a bijective homomorphism will be called an *isomorphism*.

When there is an isomorphism between semigroups S and T, we will say that S and T are *isomorphic*, and write $S \cong T$. ■

Examples 1.2.23. • Let $\varphi : \mathbb{N} \to \mathbb{N}^0$ be given by $\varphi(n) = n$ for all $n \in \mathbb{N}$ (the natural inclusion map). Then φ is a semigroup homomorphism, which is injective but not surjective, as there is no $n \in \mathbb{N}$ such that $\varphi(n) = 0$;

• Conversely, there are no homomorphisms from $\mathbb{N}^0$ to $\mathbb{N}$, for any such homomorphism φ would have to satisfy

$$\varphi(0) = \varphi(0 + 0) = 2\varphi(0),$$

which is impossible for $\varphi(0)$ a natural number;

• Let S be the semigroup $(\{a, b\}, \star)$ with multiplication table

$\star$	a	b
a	a	a
b	b	b

.

Define $\pi : S \to \{0, 1\}$ by $\pi(a) = 0$, and $\pi(b) = 1$. Then π is a bijection, but as

$$\pi(ba) = \pi(b) = 1,\ \pi(b)\pi(a) = 0 \cdot 1 = 0,$$

then π is not an isomorphism, as $\pi(ba) \neq \pi(b)\pi(a)$. △

A semigroup homomorphism φ between monoids M and N need not be a monoid homomorphism, as φ does not have to map 1_M to 1_N. However, the following lemma shows that semigroup epimorphisms between monoids are also monoid homomorphisms.

Lemma 1.2.24. *Let M and N be monoids, and let $\varphi : M \to N$ be a semi-*

group homomorphism. If φ is surjective, then φ is a monoid homomorphism.

Proof. For any $n \in N$, there exists some $m \in M$ such that $\varphi(m) = n$ by surjectivity of φ. Then

$$\varphi(1_M)n = \varphi(1_M)\varphi(m) = \varphi(1_M m) = \varphi(m) = n,$$

and a similar proof shows that $n\varphi(1_M) = n$. Hence we have shown that $\varphi(1_M)$ is an identity for N.

As the identity of N is unique, it must be that $\varphi(1_M) = 1_N$, and hence φ is a monoid homomorphism. □

1.3 Relations, congruences, quotients of semigroups & monoids

Given semigroup elements s and t, we can ask if they are somehow abstractly *related* to each other. The study of particular relations on semigroups called *congruences* are of particular interest, which we will define in this section. We first remind the reader of *equivalence relations* in the following definition.

Definition 1.3.1. A *binary relation* on a set X is a subset of $X \times X$. An *equivalence relation* $\sim$ on a set X is a subset $\sim \subseteq X \times X$ satisfying the following properties:

(R) Reflexivity: For all $x \in X$, $(x, x) \in \sim$;

(S) Symmetry: If $(x, y) \in \sim$, then $(y, x) \in \sim$;

(T) Transitivity: If $(x, y) \in \sim$ and $(y, z) \in \sim$, then $(x, z) \in \sim$.

The $\sim$-*equivalence class* of an element $x \in X$ is the set

$$[x]_\sim = \{y \in X : (x, y) \in \sim\}.$$

The *quotient set of S by* $\sim$ is the set $X/\sim$ of all $\sim$-equivalence classes of X,

given by

$$X/\sim = \{[x]_\sim \in \mathcal{P}(X) : x \in X\}. \qquad \blacksquare$$

We will often write $x \sim y$ to mean $(x, y) \in \sim$, and say that x is "$\sim$-related" to y. This motivates the occasional use of notation such as "$=$" and "$\equiv$" as subsets.

Examples 1.3.2. • The subset $\Delta = \{(x, x) : x \in X\}$ of $X \times X$ is an equivalence relation on any set X, called the *equality relation on* X;

• The subset $\nabla = X \times X$ is an equivalence relation on any set X, called the *universal relation on* X;

• The subset $\equiv_2 = \{(x, y) \in \mathbb{N} \times \mathbb{N} : x \equiv y \pmod 2\}$ is an equivalence relation on $\mathbb{N}$;

• If X is a set of people, then the subset

$$\sim = \{(x, y) \in X \times X : x \text{ and } y \text{ have the same birthday}\}$$

is an equivalence relation on that set of people. $\triangle$

Given two equivalence relations σ and ρ on the same set X, we can define their composition as follows.

Definition 1.3.3. Let σ, ρ be two equivalence relations on a set X. Then the *composition of* σ and ρ is the relation denoted $\sigma \circ \rho$ on X, defined by

$$(x, y) \in \sigma \circ \rho \Leftrightarrow \exists\, z \in X \text{ such that } (x, z) \in \sigma, (z, y) \in \rho. \qquad \blacksquare$$

Lemma 1.3.4. *Let σ, ρ be two equivalence relations on a set X. If $\sigma \circ \rho = \rho \circ \sigma$, then $\sigma \circ \rho$ is an equivalence relation on X.*

Proof. Firstly, as σ and ρ are reflexive relations, then $(x, x) \in \sigma$ and $(x, x) \in \rho$ for all $x \in X$, and hence $(x, x) \in \sigma \circ \rho$, so that $\sigma \circ \rho$ is reflexive.

Secondly, suppose $(x, y) \in \sigma \circ \rho$. Then there exists some $z \in X$ such that $(x, z) \in \sigma$ and $(z, y) \in \rho$. As σ, ρ are symmetric, then $(y, z) \in \rho$, and

$(z, x) \in \sigma$. Hence $(y, x) \in \rho \circ \sigma = \sigma \circ \rho$, and $\sigma \circ \rho$ is symmetric.

Finally, suppose that $(x, y) \in \sigma \circ \rho$, and $(y, z) \in \sigma \circ \rho$. Then there exist $s, t \in X$ such that

$$(x, s) \in \sigma,\ (s, y) \in \rho,\ (y, t) \in \sigma,\ (t, z) \in \rho.$$

In particular, $(s, t) \in \rho \circ \sigma = \sigma \circ \rho$ by definition, hence there exists some $u \in X$ such that $(s, u) \in \sigma$, $(u, t) \in \rho$.

It now follows that $(x, u) \in \sigma$ by transitivity, and $(u, z) \in \rho$ again by transitivity. Hence $(x, z) \in \sigma \circ \rho$, and $\sigma \circ \rho$ is transitive, giving the result. □

Equivalence relations on semigroups that also respect the multiplication of the semigroup are of special concern, and will be called *congruences* in the following definition.

Definition 1.3.5. For a semigroup S, a *congruence* on S is an equivalence relation σ on S such that

$$(s, t) \in \sigma \text{ and } (u, v) \in \sigma \Rightarrow (su, tv) \in \sigma.$$

In this way, σ is said to be *compatible* with the multiplication of S. ■

Examples 1.3.6. • The equality relation $\Delta = \{(s, s) \in S \times S\}$ from Examples 1.3.2 is an equivalence relation on any semigroup S, and moreover if $(s, s) \in \Delta$, and $(t, t) \in \Delta$, then $(st, st) \in \Delta$. Hence the equality relation is a congruence on any semigroup S;

• Similarly, the universal relation $\nabla = S \times S$ is a congruence on any semigroup S, as $(s, t) \in \nabla$ and $(u, v) \in \nabla$ imply that $s, t, u, v \in S$, and hence $su, tv \in S$, so that $(st, uv) \in \nabla$;

• The equivalence relation $\equiv_2 = \{(x, y) \in \mathbb{N} \times \mathbb{N} : x \equiv y \,(\mathrm{mod}\, 2)\}$ is a congruence, as if $(x, y), (m, n) \in 2\mathbb{N}$, then $x \equiv y \,(\mathrm{mod}\, 2)$, $z \equiv t \,(\mathrm{mod}\, 2)$, and hence $xz \equiv yt \,(\mathrm{mod}\, 2)$. Hence $(xz, yt) \in \equiv_2$. △

We next note a sufficient condition for the composition of two congruences

to be a congruence.

Lemma 1.3.7. *Let σ, ρ be two congruences on a semigroup S. If $\sigma \circ \rho = \rho \circ \sigma$, then $\sigma \circ \rho$ is a congruence on S.*

Proof. That $\sigma \circ \rho$ is an equivalence relation is given by Lemma 1.3.4. Let $(s,t) \in \sigma \circ \rho$ and $(u,v) \in \sigma \circ \rho$. Then there exist $x, y \in S$ such that

$$(s,x) \in \sigma.\ (x,t) \in \rho,\ (u,y) \in \sigma,\ (y,v) \in \rho.$$

As σ and ρ are congruences on S, then it follows that $(su, xy) \in \sigma$, and $(xy, tv) \in \rho$. Hence $(su, tv) \in \sigma \circ \rho$, and thus $\sigma \circ \rho$ is a congruence on S. □

For a semigroup S, any subset X of $S \times S$ is contained in some congruence on S ($S \times S$ is itself a congruence). Hence we give the following definition.

Definition 1.3.8. For a binary relation R on S, the *congruence generated by R* is the smallest congruence on S containing R (or equivalently, the intersection of all congruences on S containing R). We will denote this congruence by $R^\sharp$. ■

Given a congruence σ on a semigroup S, we are able to give the quotient set S/σ an algebraic structure, similarly to how quotients of groups by normal subgroups are obtained.

Lemma 1.3.9. *Let S be a semigroup, and let σ be a congruence on S. Then the quotient set*

$$S/\sigma = \{[s]_\sigma : s \in S\},$$

together with the operation $\cdot : S/\sigma \to S/\sigma$ defined

$$[s]_\sigma \cdot [t]_\sigma = [st]_\sigma$$

is a semigroup. □

Definition 1.3.10. For a semigroup S and a congruence σ on S, the quotient

set S/σ together with the operation $\cdot : S/\sigma \to S/\sigma$ defined

$$[s]_\sigma \cdot [t]_\sigma = [st]_\sigma$$

is called the *quotient semigroup of S by σ*. ■

For a monoid M with identity 1, and for a congruence σ on M, the quotient semigroup M/σ is actually a monoid. The identity is $[1]_\sigma$, as

$$[1]_\sigma \cdot [m]_\sigma = [1m]_\sigma = [m]_\sigma$$

and similarly $[m]_\sigma \cdot [1]_\sigma = [m]_\sigma$ for all $m \in M$. Hence we will refer to M/σ as the *quotient monoid of M by σ*.

Having defined quotient semigroups, there is a natural semigroup analogy to the first isomorphism theorem for groups, which we state in the following result.

Theorem 1.3.11 (First isomorphism theorem)**.** *Let S and T be semigroups, and let $\varphi : S \to T$ be a semigroup homomorphism. Then*

$$\ker \varphi := \{(s,t) \in S \times S : \varphi(s) = \varphi(t)\}$$

is a congruence on S,

$$\operatorname{im} \varphi := \{t \in T : t = \varphi(s) \text{ for some } s \in S\}$$

is a subsemigroup of T, and

$$S/\ker \varphi \cong \operatorname{im} \varphi.$$ □

1.4 Green's relations on semigroups and monoids

Particularly notable examples of relations on a semigroup are *Green's relations*, which give us a way of understanding the ideal substructures of a semigroup. We now define Green's relations $\mathcal{L}, \mathcal{R}, \mathcal{H}$ and $\mathcal{J}$ on any semigroup S.

Definition 1.4.1. Recall the definitions of the principal ideals generated by an element a of a semigroup S from Definition 1.2.20 and Definition 1.2.19. Then *Green's relations* $\mathcal{L}, \mathcal{R}, \mathcal{H}$, and $\mathcal{J}$ on a semigroup S are the binary relations defined as follows.

$$\begin{aligned}(a,b) \in \mathcal{L} &\Leftrightarrow S^1a = S^1b,\\ (a,b) \in \mathcal{R} &\Leftrightarrow aS^1 = bS^1,\\ (a,b) \in \mathcal{H} &\Leftrightarrow (a,b) \in \mathcal{L} \cap \mathcal{R},\\ (a,b) \in \mathcal{J} &\Leftrightarrow S^1aS^1 = S^1bS^1.\end{aligned}$$

It is a brief exercise to show that $\mathcal{L}, \mathcal{R}, \mathcal{H}$ and $\mathcal{J}$ are equivalence relations on S. Hence the $\mathcal{L}, \mathcal{R}, \mathcal{H}$ and $\mathcal{J}$-classes of an element $a \in S$ will be denoted by L_a, R_a, H_a and J_a, respectively. Note that

$$H_a = L_a \cap R_a,\ H_a \subseteq L_a,\ H_a \subseteq R_a,\ L_a \subseteq J_a,\ R_a \subseteq J_a. \qquad \blacksquare$$

Definition 1.4.2. For $\mathcal{K}$ any of Green's relations $\mathcal{L}, \mathcal{R}, \mathcal{H}$ or $\mathcal{J}$, a semigroup S is said to be *$\mathcal{K}$-trivial* if $\mathcal{K}$ is equal to the equality relation Δ. $\blacksquare$

Examples 1.4.3. • For $a \in \mathbb{N}$, we saw in Examples 1.2.21 that the principal left, right, and two sided ideals of $(\mathbb{N}, +)$ generated by a were given by

$$S^1a = aS^1 = S^1aS^1 = \{x \in \mathbb{N} : x \geq a\}.$$

Denoting the set $\{x \in \mathbb{N} : x \geq a\}$ by I_a, we note that $I_a = I_b \Leftrightarrow a = b$. That is, $(a,b) \in \mathcal{K} \Rightarrow a = b$, where $\mathcal{K}$ is any of Green's relations $\mathcal{L}, \mathcal{R}, \mathcal{H}$ or $\mathcal{J}$.

Hence it follows that $\mathcal{L}, \mathcal{R}, \mathcal{H}$ and $\mathcal{J}$ on S are all equal to the equality relation

$$\Delta = \{(s,s) : s \in S\},$$

and so $(\mathbb{N}, +)$ is $\mathcal{L}$-trivial, $\mathcal{R}$-trivial, $\mathcal{H}$-trivial and $\mathcal{J}$-trivial.

• For $(\mathbb{N}, \star)$ given in Examples 1.2.13, we saw in Examples 1.2.21 that for all $a \in \mathbb{N}$, the principal left ideal S^1a of $S = \mathbb{N}$ is equal to $\mathbb{N}$. Hence for all

$a, b \in S$, $S^1a = S^1b$, and so $\mathcal{L}$ is the universal relation

$$\nabla = S \times S.$$

We also saw that for $a \in S$, every principal two-sided ideal S^1aS^1 of $\mathbb{N}$ is also equal to $\mathbb{N}$, and hence similarly to $\mathcal{L}$, we find that Green's relation $\mathcal{J}$ on $(\mathbb{N}, \star)$ is the universal relation ∇.

Finally, we saw that for $a \in S$, every principal right ideal aS^1 was equal to the singleton set $\{a\}$. Hence $aS^1 = bS^1$ if and only if $a = b$, and so $\mathcal{R}$ is the equality relation

$$\Delta = \{(s, s) : s \in S\}.$$

As $\mathcal{H} \subseteq \mathcal{R}$, it follows that Green's relation $\mathcal{H}$ on $(\mathbb{N}, \star)$ is also the equality relation Δ. Thus $(\mathbb{N}, \star)$ is $\mathcal{R}$-trivial and $\mathcal{H}$-trivial. $\triangle$

The set of $\mathcal{L}$-classes of S correspond to the principal left ideals of S, the set of $\mathcal{R}$-classes correspond to the principal right ideals, and the $\mathcal{J}$-classes correspond to the principal two-sided ideals. In this way, Green's relations tell us about the ideal substructures contained within the semigroup S, hence their importance.

We now derive some facts about these relations that will be needed for later results of this thesis.

We begin by giving an alternate but equivalent way of viewing $\mathcal{L}, \mathcal{R}$ and $\mathcal{J}$ in terms of the multiplication of a semigroup S in the following lemma.

Lemma 1.4.4. *Let S be a semigroup, and let $a, b \in S$. Then*

(i) $(a, b) \in \mathcal{L} \Leftrightarrow (\exists x, y \in S^1)(a = xb)(b = ya)$;

(ii) $(a, b) \in \mathcal{R} \Leftrightarrow (\exists x, y \in S^1)(a = bx)(b = ay)$;

(iii) $(a, b) \in \mathcal{J} \Leftrightarrow (\exists x, y, u, v \in S^1)(a = xby)(b = uav)$.

Proof. (i) ($\Rightarrow$) Let $(a, b) \in \mathcal{L}$, so that $S^1a = S^1b$. As $a = 1a \in S^1a$, then $a \in S^1b$. Hence there exists some $x \in S^1$ such that $a = xb$. Similarly, as $b = 1b \in S^1b$, then $b \in S^1a$ and so there exists some $y \in S^1$ such that $b = ya$.

($\Leftarrow$) Let $a, b \in S$, and let $x, y \in S^1$ be such that $a = xb$ and $b = ya$. Then as $x \in S^1$, it follows that $a = xb \in S^1 b$. Letting $s \in S^1$, then

$$sa = s(xb) = (sx)b \in S^1 b,$$

and hence as s was arbitrarily chosen, we have shown that $S^1 a \subseteq S^1 b$. Conversely, as $y \in S^1$, then $b = ya \in S^1 a$. Again, letting $s \in S^1$, then

$$sb = s(ya) = (sy)a \in S^1 a,$$

and thus $S^1 b \subseteq S^1 a$. Hence $S^1 a = S^1 b$, and so $(a, b) \in \mathcal{L}$.

(ii) The proof is similar to that of (i) by left/right symmetry.

(iii) ($\Rightarrow$) Let $(a, b) \in \mathcal{J}$, so that $S^1 a S^1 = S^1 b S^1$. Then $a = 1a1 \in S^1 a S^1$, and hence $a \in S^1 b S^1$. Thus there exists some $x, y \in S^1$ such that $a = xby$. The proof that $b = uav$ for some $u, v \in S^1$ is similar.

($\Leftarrow$) For $a, b \in S$, suppose there exists $x, y, u, v \in S^1$ satisfying $a = xby$, $b = uav$. Then for $s, t \in S^1$, we have

$$sat = s(xby)t = (sx)b(yt) \in S^1 b S^1,$$

and hence as s, t were arbitrarily chosen, then $S^1 a S^1 \subseteq S^1 b S^1$. The proof for the converse, that $S^1 b S^1 \subseteq S^1 a S^1$ is similar, and hence $S^1 a S^1 = S^1 b S^1$, meaning $(a, b) \in \mathcal{J}$. □

The following result, commonly known as *Green's lemma*, gives two mutually inverse bijections between the $\mathcal{R}$-classes of two $\mathcal{L}$-related elements of a semigroup S. The proof is omitted, as a comprehensive version is given by Howie [16, Lemma 2.2.1].

Lemma 1.4.5 (Green's lemma [16, Lemma 2.2.1])**.** *Let S be a semigroup, let $a, b \in S$ be such that $(a, b) \in \mathcal{L}$, and let $x, y \in S^1$ be such that*

$$a = xb,\ b = ya.$$

For $t \in S$, define the left translation map λ_t by

$$\lambda_t : S \to S := \lambda_t(s) = ts.$$

Then the restriction $\lambda_y|_{R_a}$ is a bijection from R_a onto R_b, and the restriction $\lambda_x|_{R_b}$ is a bijection from R_b onto R_a, as

$$\lambda_y|_{R_a} \circ \lambda_x|_{R_b} = \mathrm{id}_{R_a},$$
$$\lambda_x|_{R_b} \circ \lambda_y|_{R_a} = \mathrm{id}_{R_b},$$

where id_{R_a} and id_{R_b} are the identity maps on R_a and R_b, respectively.

Moreover, if $u \in R_a$, then $(u, (\lambda_y|_{R_a})(u)) \in \mathcal{L}$, and if $v \in R_b$, then $(v, (\lambda_x|_{R_b})(v)) \in \mathcal{L}$. □

There is a natural dual to Green's lemma (which is also referred to as "Green's lemma"), which gives two mutually inverse bijections between the $\mathcal{L}$-classes of two $\mathcal{R}$-related elements S (see [16, Lemma 2.2.2]).

Green's lemmas are important structural results in their own right: they tell us that the $\mathcal{L}$ (or $\mathcal{R}$) classes of $\mathcal{R}$ (or $\mathcal{L}$) related elements are of the same cardinality. They are also used in the proof of *Green's theorem*, which gives us a characterisation of when an $\mathcal{H}$-class of a semigroup is actually a subgroup. We now state this theorem without proof below, again referring the reader to Howie [16, Theorem 2.2.5].

Theorem 1.4.6 (Green's theorem [16, Theorem 2.2.5])**.** *Let S be a semigroup, and let H be a $\mathcal{H}$-class of S. Then either $H^2 \cap H = \emptyset$, or $H^2 = H$ and H is a subgroup of S.* □

We obtain the following useful result, as a corollary of Green's theorem, which relates the idempotents of a semigroup to its subgroups.

Corollary 1.4.7. *If S is a semigroup, and $e \in E(S)$ is an idempotent, then H_e is a subgroup of S.*

Proof. As e is idempotent, then $e^2 = e$, and hence $H_e^2 \cap H_e \neq \emptyset$. Hence by

Green's theorem, H_e is a subgroup of S. □

1.5 Formal language theory, automata, free semigroups & monoids

An area of overlap with that of semigroup theory and universal algebra is the theory of *formal languages* in theoretical computer science. Many problems in semigroup theory can be formulated as algorithmic problems of *decision*, such as the word problem in group theory for example. That is, given a finite set of information, one asks whether there exists an algorithm taking this information, and deciding a yes or no output for a given question.

Very simple such algorithms can be given with the machinery of *automata*, which (in general) take a formal *word* and either accept or reject it. The set of these accepted words form a *language*, which can be described (as we will define shortly) as a subset of a *free monoid* over a given *alphabet*.

Free semigroups and monoids in particular will be ubiquitous in this thesis. Moreover, we will later be considering decision problems for semigroups, such as the *word problem*, *membership problem* and the *finite generation problem* (defined in Section 1.6). Hence in this section, we will introduce the nomenclature just described, as used in this work. For a more comprehensive study on these concepts, we refer the reader to [24].

We begin in the theory of lanuages, with the following definition.

Definition 1.5.1. An *alphabet* is simply a set $A = \{a_1, a_2, \dots\}$, typically consisting of symbols and characters (though A can be any set). A *letter* is an element of the alphabet A.

A *word w over the alphabet A* is a finite sequence of letters of A, typically written consecutively in the format

$$w = a_{i_1} a_{i_2} a_{i_3} \dots a_{i_m}.$$

When the context of A is clear, we will simply refer to w as a word. The *i-th*

letter of w is the i-th element of the sequence describing w. It is denoted by $|w|_i$.

If $a \in A$, and w is a word over A, then the *a-count* of w is the number of times a occurs in the sequence describing w. This is denoted by $|w|_a$.

The *length* of a word w is the number of letters in it, or equivalently, the length of the sequence describing it. The length of a word w will be denoted $|w|$.

The *empty word* ε_A over A is the empty sequence of letters, or equivalently, the word over A consisting of no letters. By convention, the empty word has length 0. When the context of A is clear, we will simply write ε to denote the empty word. ■

By convention, if a letter $a \in A$ appears n times consecutively in a word w over A, we will often write a^n rather than

$$\underbrace{aaa\ldots a}_{n \text{ times}}.$$

For example, the word $baab$ over the alphabet $A = \{a, b\}$ would be written as ba^2b.

Definition 1.5.2. The set of all non-empty words over A will be denoted by A^+, and the set of all words over A will be denoted by A^*. Note that $A^* = A^+ \cup \{\varepsilon\}$. The *concatenation* of two words $u = a_1a_2\ldots a_n \in A^+$ and $v = b_1b_2\ldots b_m \in A^+$ (where $a_1, \ldots, a_m, b_1, \ldots, b_n \in A$) is the word $w = uv$ over A, where uv is the concatenation of u and v as sequences. That is,

$$w = a_1a_2\ldots a_nb_1b_2\ldots b_m.$$

The set A^+ with the operation of concatenation of words is verifiably a semigroup, and is hence called the *free semigroup* on A. If A is a finite set, the free semigroup A^+ will be said to have *rank* $|A|$.

Similarly, A^* with the operation of concatenation is a monoid with identity

ε, when adopting the convention that

$$\varepsilon w = w\varepsilon = w$$

for all $w \in A^*$. A^* will be called the *free monoid* over A. ■

Definition 1.5.3. Let A be an alphabet, and let $w \in A^*$ be a word. Then a *prefix* of w is a word $u \in A^*$ such that there exists $v \in A^*$ with

$$w = uv,$$

where uv is the concatenation of u with v. Note that the empty word and the word w itself are always prefixes of any word w. We will use the notation $u \leq_p w$ to mean that u is a prefix of w.

Similarly, a *suffix* of w is a word $v \in A^*$ such that there exists $u \in A^*$ with

$$w = uv.$$

Again, note that the empty word and the word w itself are always suffixes of any word w. We will use the notation $v \leq_s w$ to mean v is a suffix of w.

A *proper* prefix u of w is a prefix u which is not equal to the whole word w itself, or the empty word. We will write $u <_p w$ in this case. Similarly, a *proper suffix* v of w is a suffix which is not equal to w or the empty word. In this case, we will write $v <_s w$. ■

Definition 1.5.4. Let u be a prefix of a word w, so that $uv = w$ for some suffix v of w. Then the *word w stripped of the prefix u* is the word v such that $uv = w$. Similarly, for a suffix v of w, the *word w stripped of the suffix v* is the word u such that $uv = w$.

We denote the word w stripped of prefix u by $u^{-1}w$. Similarly, we denote the word w stripped of suffix v by wv^{-1}. ■

Examples 1.5.5. • 01001110_01001001_01001011 is a word over the alphabet $\{0, 1, \_\}$, of length 26;

• $ABBA$ is a word of length 4 over the two letter alphabet $\{A, B\}$, $Waterloo$ is a word of length 8 over the alphabet $\{a, e, l, o, r, t, W\}$. Both are words over the alphabet $\{a, e, l, o, r, t, W, A, B\}$. The W-count of $Waterloo$ is 1, but the o-count of $Waterloo$ is 2.

• $Water$ is a prefix of $Waterloo$, whereas loo is a suffix. Both are proper.

• $(Water)^{-1}(Waterloo) = loo$, and $(Waterloo)(loo)^{-1} = Water$. △

1.6 Generating and presenting semigroups and monoids

Of course, not every subset of a semigroup is a subsemigroup. However, for a semigroup S and a subset $X \subseteq S$ which is non-empty, there is always a subsemigroup of S containing X (S itself would do). In this section, we discuss how to generate subsemigroups from a given subset of a semigroup. We will discuss related concepts such as indecomposability for semigroups, finite generation, and finish with the theory of semigroup and monoid presentations.

To begin, we have already noted that for any non-empty subset X of a semigroup S, there exists a subsemigroup of S containing X. Hence the intersection of all subsemigroups of S that contain X is non-empty, and it is a short exercise to see that this intersection is also a subsemigroup of S. Hence we make the following definition.

Definition 1.6.1. For a semigroup S and a subset $X \subseteq S$, the *subsemigroup generated by* X is the intersection of all subsemigroups of S containing X. The subsemigroup generated by X will be denoted $\langle X \rangle$, and X will be called a *semigroup generating set for* $\langle X \rangle$.

Equivalently, $\langle X \rangle$ is the smallest subsemigroup of S containing X, and can equivalently be defined as the set of all finite products of elements in X. The elements of X will be called the *generators* of $\langle X \rangle$. ■

By convention, we will write $\langle x \rangle$ instead of $\langle \{x\} \rangle$, when $X = \{x\}$ is a singleton set. We note that every instance of the word *semigroup* can be replaced in

the above definition by *monoid* to obtain the definition for the *submonoid generated by* X.

In this instance, the identity element 1 is not required to be in a monoid generating set X for $\langle X \rangle$, as $\langle X \rangle$ is a submonoid and hence contains 1.

Examples 1.6.2. • Let $X = \{1\}$ be the singleton subset of $\mathbb{N}$. Then as $1 \in \langle X \rangle$ and $\langle X \rangle$ is a subsemigroup of $(\mathbb{N}, +)$, it must follow that

$$n = \underbrace{1 + 1 + \cdots + 1}_{n \text{ times}} \in \langle X \rangle$$

for all $n \in \mathbb{N}$. Hence $\mathbb{N} \subseteq \langle X \rangle$, and it follows that $\mathbb{N} = \langle X \rangle$. Hence $\{1\}$ is a semigroup generating set for $\mathbb{N}$;

• By the same reasoning, $\{1\}$ is a monoid generating set for $(\mathbb{N}^0, +)$;

• Let X be any non-empty subset of $\mathbb{N}$, and consider the semigroup $(\mathbb{N}, \star)$, where $\star : \mathbb{N} \to \mathbb{N}$ is defined by $m \star n = m$ for all $m, n \in \mathbb{N}$. As $\langle X \rangle$ can be considered as the set of all finite products of elements of X, then as

$$n_1 \star n_2 \star \cdots \star n_k = n_1$$

for any finite product of elements $n_1, n_2, \ldots, n_k \in X$, it follows that $\langle X \rangle \subseteq X$. As $X \subseteq \langle X \rangle$ by definition, it follows that $X = \langle X \rangle$, and hence in particular any non-empty subset X of $\mathbb{N}$ is a subsemigroup of $(\mathbb{N}, \star)$. △

We now give a definition of the *order* of a semigroup element, noting that this definition generalises the concept of the order of a group element.

Definition 1.6.3. An element x of a semigroup S is said to have *finite order* if $\langle x \rangle$ is finite. Otherwise, x is said to have *infinite order*.

The *order* of an element x (of finite order) is the size of $\langle x \rangle$. ■

In the next definition, we distinguish the specific case where there exists a singleton generating set for a semigroup (such as $\mathbb{N}$ for example).

Definition 1.6.4. Let S be a semigroup. If there exists $x \in S$ such that

$$\langle x \rangle = \{x, x^2, x^3, \dots\} = S,$$

then S will be called a *monogenic semigroup*. ■

Given a generating set X for a semigroup S, and an element $s \in S$, then s can certainly be written as a finite product of the elements of X. If $s = xy$, then we will say that s can be *decomposed* into the product of x and y.

However not all elements will have this property (for example, $1 \in \mathbb{N}$ cannot be written as the sum of two natural numbers). Hence we give the following definition to distinguish these elements.

Definition 1.6.5. Let S be a semigroup, let $X \subseteq S$ be any non-empty subset, and let $s \in S$. If there exist two elements $x, y \in \langle X \rangle$ such that

$$s = xy,$$

then s will be called *decomposable over* X. That is, $s \in \langle X \rangle^2$.

Otherwise, s is said to be *indecomposable over* X. When $X = S$, we will also say that s is either a *(semigroup) decomposable* or *(semigroup) indecomposable* element of S, respectively. ■

Of course, if S is a monoid, then any element of S is a semigroup decomposable element, but it might be that a given element is only trivially decomposable into a product involving the identity. We thus distinguish monoid indecomposability in the following definition.

Definition 1.6.6. Let M be a monoid with identity 1, and let $m \in M$. If there exist $x, y \in M \setminus \{1\}$, such that

$$m = xy,$$

then m will be called a *(monoid) decomposable* element of M. Otherwise, m will be called a *(monoid) indecomposable* element of M. ■

Examples 1.6.7. • In $(\mathbb{N}, +)$, the only indecomposable element is 1, as any other element $n \in \mathbb{N}$ can be written as the sum of 1 and $n - 1$;

• In the monoid $(\mathbb{N}^0, +)$, every element is decomposable in the semigroup sense, but 1 is the only indecomposable element in the monoid sense for the same reasoning as above;

• In the monoid $(\{0, 1\}, \times)$, 0 is decomposable in both the semigroup and monoid sense as $0 = 0^2$, but 1 is decomposable in the semigroup sense and indecomposable in the monoid sense, as $1 = 1^2$. △

We will next see that given a generating set X for a semigroup S, that X must contain every semigroup indecomposable element of S.

Lemma 1.6.8. *If X is any semigroup generating set for a semigroup S, then X contains the set of semigroup indecomposable elements of S.*

Proof. Let $s \in S$ be a semigroup indecomposable element of S. Then as $S = \langle X \rangle$, s can be written as a finite product of elements of X. There cannot exist a decomposition

$$s = x_1 x_2 \dots x_n$$

of s into a product of n elements of X where $n \geq 2$, as this contradicts indecomposability of S.

Hence it must be that s is an element of X, and hence the set of all indecomposable elements of S is contained within X. □

We now will see that images of generating sets under epimorphisms are also generating sets, as in the following pair of lemmas.

Lemma 1.6.9. *Let M, N be monoids, let $X \subseteq M$ be a monoid generating set for M, and let $\varphi : M \to N$ be a monoid epimorphism. Then $\varphi(X)$ is a monoid generating set for N.*

Proof. Firstly, as φ is surjective, then every $n \in N \setminus \{1_N\}$ has some

$m \in M \setminus \{1_M\}$ such that $n = \varphi(m)$. Moreover, as any $m \in M \setminus \{1_M\}$ can be written as a finite product $m = x_1 x_2 \dots x_p$ of letters x_i from X, then it follows that every $n \in N \setminus \{1_N\}$ can be written as a finite product $n = \varphi(x_1)\varphi(x_2)\dots\varphi(x_p)$ of images $\varphi(x_i)$ of letters x_i from X. Hence $\varphi(X)$ is a monoid generating set for N. □

Lemma 1.6.10. *Let S, T be semigroups, let $X \subseteq S$ be a semigroup generating set for S, and let $\varphi : S \to T$ be a semigroup epimorphism. Then $\varphi(X)$ is a semigroup generating set for T.*

Proof. The proof is the same as for Lemma 1.6.9, replacing instances of the word monoid for semigroup, and ignoring instances of identities. □

We next note that indecomposable elements of a semigroup S are preserved under isomorphism.

Lemma 1.6.11. *Let S, T be isomorphic semigroups, let $\varphi : S \to T$ be a semigroup isomorphism between them, and let $\mathcal{I}(S)$ and $\mathcal{I}(T)$ be the set of semigroup indecomposable elements of S and T respectively. Then*

$$\varphi(\mathcal{I}(S)) = \mathcal{I}(T).$$

Proof. Let $s \in S$ be a semigroup indecomposable element of S. Suppose for a contradiction that $\varphi(s)$ is a semigroup decomposable element of T. Then there exist $x, y \in T$ such that

$$\varphi(s) = xy.$$

As φ is an isomorphism, then it is surjective in particular. Hence there exist $u, v \in S$ such that $x = \varphi(u)$ and $y = \varphi(v)$, and hence

$$\varphi(s) = \varphi(u)\varphi(v) = \varphi(uv).$$

As φ is injective, then $s = uv$. This is a contradiction, as now s is semigroup decomposable in S. Thus $\varphi(s) \in \mathcal{I}(T)$, and hence as $s \in \mathcal{I}(S)$ was chosen

arbitrarily, we have $\varphi(\mathcal{I}(S)) \subseteq \mathcal{I}(T)$. The reverse inclusion follows from the same argument by applying the inverse of φ, and hence the result follows. □

We now return to concepts relating to generating sets for semigroups. As a group may be described as *finitely generated*, so too can a semigroup or a monoid, as we see in the following definition.

Definition 1.6.12. A semigroup S is said to *finitely generated as a semigroup* if there is a finite subset $X \subseteq S$ such that $S = \langle X \rangle$. In this instance, S will be called a *finitely generated semigroup*.

Similarly, a monoid M is said to be *finitely generated as a monoid* if there is a finite subset $X \subseteq M$ such that $M = \langle X \rangle$, where $\langle X \rangle$ is the *submonoid* generated by X. ■

Note that finitely generated groups are finitely generated as monoids, and finitely generated monoids are finitely generated as semigroups.

Examples 1.6.13. • $(\mathbb{N}, +)$ is finitely generated as a semigroup by the set $\{1\}$, as we saw in Examples 1.6.2;

• Similarly, $(\mathbb{N}^0, +)$ is finitely generated as a monoid, again by $\{1\}$;

• $(\mathbb{N}, \star)$ from Examples 1.6.2 is *not* finitely generated as a semigroup, as $X = \langle X \rangle$ for all non-empty $X \subseteq \mathbb{N}$. Hence to satisfy $\langle X \rangle = \mathbb{N}$, it must be that $X = \mathbb{N}$, which is infinite. △

Every finite semigroup and monoid is of course finitely generated, by taking the whole semigroup or monoid itself as a generating set. Any property which holds for any finite semigroup, such as being finitely generated, will be called a *finitary property*.

One such related finitary property is that of being *finitely presented*, which is commonly studied in group theory. We now introduce the theory of semigroup presentations in order to define finitely presented semigroups.

Recall Definition 1.5.2, which defined the free semigroup and free monoid over a set A. A more abstract definition of what it means for a semigroup

to be *free* can be given, as in the following.

Definition 1.6.14. Given a set A, a semigroup S is said to be *free over* A if there exists a mapping $\alpha : A \to S$, such that for any semigroup T and mapping $\varphi : A \to T$, there exists a unique semigroup homomorphism $\psi : S \to T$ such that $\varphi = \alpha \circ \psi$. That is, the following diagram commutes.

$$\begin{array}{ccc} A & \xrightarrow{\alpha} & S \\ & \searrow^{\varphi} & \downarrow{\scriptstyle \psi} \\ & & T \end{array}$$

■

Of course, replacing the word semigroup in the above definition by monoid gives an abstract definition of what it means for a monoid to be free. Similarly, replacing all instances of the word semigroup by commutative semigroup above gives the definition of what it means for a commutative semigroup to be "free commutative".

We now qualify that the free semigroups and free monoids defined in Definition 1.5.2 are indeed free over their respective sets.

Lemma 1.6.15. *For any alphabet A, the free semigroup A^+ is free over A. Similarly, the free monoid A^* is free over A.*

Proof. Let A be a set. Take the mapping $\alpha : A \to A^+$ as the inclusion mapping, defined by $\alpha(a) = a$ for all $a \in A$. Suppose that φ is any mapping from A to any semigroup T. Define $\psi : A^+ \to T$ by

$$\psi(a_1 a_2 \dots a_n) = \varphi(a_1)\varphi(a_2)\dots\varphi(a_m).$$

Then ψ is a semigroup homomorphism, for if $u = a_1 a_2 \dots a_m \in A^+$ and $v = b_1 b_2 \dots b_n \in A^+$, then

$$\begin{aligned} \psi(uv) &= \psi(a_1 a_2 \dots a_m b_1 b_2 \dots b_n) \\ &= \varphi(a_1)\varphi(a_2)\dots\varphi(a_m)\varphi(b_1)\varphi(b_2)\dots\varphi(b_n) \\ &= \psi(a_1 a_2 \dots a_m)\psi(b_1 b_2 \dots b_n) \end{aligned}$$

$$= \psi(u)\psi(v).$$

Letting $a \in A$ be any element, then

$$\varphi(a) = \psi(a) = \psi(\alpha(a)) = (\alpha \circ \psi)(a),$$

and hence $\varphi = \alpha \circ \psi$. If $\psi' : A^+ \to T$ is any other semigroup homomorphism such that $\varphi = \alpha \circ \psi'$, then $\alpha \circ \psi' = \alpha \circ \psi$. But for all $a \in A$, as

$$(\alpha \circ \psi)(a) = \psi(\alpha(a)) = \psi(a),$$

then $\alpha \circ \psi = \psi$. Similarly, $\alpha \circ \psi' = \psi'$, and hence it follows that $\psi = \psi'$. Thus ψ is the unique semigroup homomorphism such that $\varphi = \alpha \circ \psi$. Hence A^+ is free over A.

The case showing the free monoid A^* is free over A is the same as A^+, taking instead T to be any monoid M with identity 1_M, and the map $\psi : A^* \to M$ to be given by

$$\psi : A^* \to M := \psi(u) = \begin{cases} \varphi(a_1)\varphi(a_2)\dots\varphi(a_m) & \text{if } u = a_1a_2\dots a_m \in A^+ \\ 1_M & \text{if } u = \varepsilon. \end{cases}$$

□

As an important consequence of the abstract definition of freeness, we obtain the following lemma.

Lemma 1.6.16. *Let S be a semigroup, let X be a set, and let $\varphi : X \to S$ be a mapping such that $\varphi(X)$ is a generating set for S. Then there exists a semigroup epimorphism $\psi : X^+ \to S$.*

Proof. Define $\psi : X^+ \to S$ by

$$\psi(x_1x_2\dots x_n) = \varphi(x_1)\varphi(x_2)\dots\varphi(x_n)$$

Then ψ is a semigroup homomorphism, for if $u = x_1x_2\dots x_m \in X^+$ and

$v = y_1 y_2 \dots y_n \in X^+$, then

$$\begin{aligned}\psi(uv) &= \psi(x_1 x_2 \dots x_m y_1 y_2 \dots y_n) \\ &= \varphi(x_1)\varphi(x_2)\dots\varphi(x_m)\varphi(y_1)\varphi(y_2)\dots\varphi(y_n) \\ &= \psi(x_1 x_2 \dots x_m)\psi(y_1 y_2 \dots y_n) \\ &= \psi(u)\psi(v).\end{aligned}$$

Further, as $\varphi(X)$ is a generating set for S, then for any $s \in S$, there exist $x_1, x_2, \dots, x_m \in X$ such that

$$s = \varphi(x_1)\varphi(x_2)\dots\varphi(x_m) = \psi(x_1 x_2 \dots x_m),$$

and hence ψ is surjective. □

We now are ready to define a *semigroup presentation*, which precisely formalises quotients of free semigroups by congruences.

Definition 1.6.17. Let S be a semigroup. If X is a set, and R is a binary relation on X^+ such that $S \cong \langle X : R \rangle$, where

$$\langle X : R \rangle = X^+ / R^\sharp,$$

then we say that $\langle X : R \rangle$ is a *semigroup presentation for* S. ■

Any semigroup S has a generating set (as $S = \langle S \rangle$), and hence by Lemma 1.6.16, we can always find a generating set X and epimorphism $\psi : X^+ \to S$. By the first isomorphism theorem (Theorem 1.3.11), S is isomorphic to the quotient $X^+/\ker\psi$. Hence every semigroup is isomorphic to the quotient of a free semigroup by a congruence.

Moreover, taking $R = \ker\psi$, as $\ker\psi$ is a congruence, then $R^\sharp = \ker\psi$. Hence every semigroup has a semigroup presentation.

Of course, wherever we have used the word semigroup, there is an analogy for monoids. Hence we also give the following definition.

Definition 1.6.18. Let M be a monoid. If X is a set, and R is a binary

relation on X^+ such that $M \cong \langle X : R \rangle$, where

$$\langle X : R \rangle = X^*/R^\sharp,$$

then we say that $\langle X : R \rangle$ is a *monoid presentation for* M. ■

By convention, if $X = \{x_1, x_2, \ldots x_m\}$, and $R = \{(u_1, v_1), (u_2, v_2), \ldots, (u_n, v_n)\}$ are finite sets, we will often write

$$\langle X : R \rangle = \langle x_1, x_2, \ldots, x_m : u_1 = v_1, u_2 = v_2, \ldots, u_n = v_n \rangle$$

for semigroup and monoid presentations, in a slight absuse of notation.

Example 1.6.19. Let $X = \{x\}$, and take $R = \emptyset$. Then R is trivially a binary relation on X^+. We claim that $R^\sharp = \Delta$, the equality relation on X^+.

Firstly, Δ is a congruence on X^+, as seen in Examples 1.3.6, and it contains $R = \emptyset$. Hence as $R^\sharp$ is the smallest congruence on X^+ containing R, then $R^\sharp \subseteq \Delta$. Conversely, as $\Delta = \{(w, w) : w \in X^+\}$, then any element $(w, w) \in \Delta$ is also an element of $R^\sharp$, as $R^\sharp$ is a reflexive relation. Hence $\Delta \subseteq R^\sharp$, and we have shown our claim. Thus

$$\langle x : \emptyset \rangle = X^+/\Delta,$$

which is isomorphic to $\{x\}^+$ by the isomorphism $\varphi : \{x\}^+ \to \{x\}^+/\Delta$ defined $\varphi(x^k) = [x^k]_\Delta$.

Moreover, $\{x\}^+$ is isomorphic to $\mathbb{N}$ under addition, by the isomorphism $\psi : \{x\}^+ \to \mathbb{N}$ defined $\psi(x^k) = k$. Hence $\mathbb{N}$ has semigroup presentation $\langle x : \emptyset \rangle$. △

We now define *finitely presented* semigroups and monoids, as follows.

Definition 1.6.20. For a semigroup S, if there exists a finite set X and a finite binary relation R on X^+ such that

$$S \cong \langle X : R \rangle,$$

then S will be called *finitely presented* as a semigroup, with *finite semigroup presentation* $\langle X : R \rangle$.

Similarly, for a monoid M, if there exists a finite set X and a finite binary relation R on X^* such that

$$M \cong \langle X : R \rangle,$$

then M will be called *finitely presented* as a monoid, with *finite monoid presentation* $\langle X : R \rangle$. ■

We finish this section by introducing *decision problems* that are related to the concepts of generation. The first decision problem we will describe is the *word problem* for a semigroup, defined below.

Definition 1.6.21. Given a semigroup S and $X \subseteq S$ a finite semigroup generating set for S, the *word problem of* S with respect to X is given by

$$\mathrm{WP}(S, X) = \{(u, v) \in X^+ \times X^+ : u =_S v\},$$

where $u =_S v$ if u and v represent the same element in S when written as words over the generating set X.

The word problem of S is said to be *decidable* with respect to X if there exists an algorithm taking S, a finite generating set $X \subseteq S$ and any $(u, v) \in X^+ \times X^+$ as inputs which determines whether or not $(u, v) \in \mathrm{WP}(S, X)$. ■

Examples 1.6.22. • The free semigroup A^+ over A has decidable word problem with respect to A, simply by taking two words in $u, v \in A^+$ and reading them. If $|u| \neq |v|$, then $u \neq v$. Otherwise, check if the first letter of u and v are equal, if not, then $u \neq v$. Otherwise, check if the second letter of u and v are equal, and so on. If all letters have been checked and are equal, then $u = v$;

• Any finite semigroup S has decidable word problem. Given any generating set X of S, and two words $u, v \in X^+$, there are only finitely many possibilities

for what u and v can equal in the semigroup S, hence an exhaustive algorithm checking all the possibilities is viable. △

A more general version of the word problem can be given, by asking if two words over a finite generating set are equal in a subsemigroup of a given semigroup. Hence we define the *membership problem* as follows.

Definition 1.6.23. Given a finitely generated semigroup S, a finitely generated subsemigroup T of S and a generating set X for S, the *membership problem of T in S* is the set of words over X which represent an element in T, when considered as a product of generators in X.

The membership problem is said to be *decidable* if there is an algorithm taking S, X and a finite subset Y of X^* generating T, which decides whether or not a word w over X represents an element in $\langle Y \rangle$. ■

The final decision problem of consideration in this thesis will be the *finite generation problem*, which asks if a given collection of semigroups or monoids are finitely generated, as in the following definition.

Definition 1.6.24. Given a collection $\mathcal{C}$ of semigroups (or monoids) S_i, each of which can be somehow described by a finite set of data, it is said to have *decidable finite generation problem* if there exists an algorithm taking the finite set of data, and determining whether or not each S_i is generated by some finite subset X_i of S_i. ■

1.7 Direct products, subdirect products and fiber products

In this section, we focus on the definitions which make up the main subject of study of this thesis, namely subdirect products and fiber products. We begin with an overview of the definitions in terms of semigroups, alongside examples. We then introduce a number of important results such as Fleischer's lemma, and its associated corollaries. Finally, we highlight some of the motivating examples of subdirect products of groups that we introduced

in this chapter, in greater specificity.

We begin by recalling the *direct product* structure for groups, which we now define for semigroups.

Definition 1.7.1. Given two semigroups S and T, the *direct product of S with T* is the Cartesian product

$$S \times T = \{(s,t) : s \in S, t \in T\},$$

together with pointwise multiplication

$$(s,t)(u,v) = (su, tv),$$

where su is a product in S, and tv is a product in T. In this way, $S \times T$ is a semigroup. S and T are called *factors* of the direct product. The maps defined by

$$\pi_S : S \times T \to S := (s,t) \mapsto s,$$
$$\pi_T : S \times T \to T := (s,t) \mapsto t.$$

are known as the *projection maps* onto S and T, respectively. ■

When S and T are both monoids or both groups, then $S \times T$ is a monoid or a group. Of course, we can also define a finite direct product of n semigroups to be the set of all n-tuples with pointwise multiplication. We outline the specific case when the factors are all the same in the following definition.

Definition 1.7.2. Let S be a semigroup. For $k \in \mathbb{N}$, the *finite direct power* S^k is the k-fold Cartesian product

$$S^k = \underbrace{S \times S \times \cdots \times S}_{k \text{ times}},$$

together with the pointwise operation

$$(s_1, s_2, \ldots s_k)(t_1, t_2, \ldots, t_k) = (s_1 t_1, s_2 t_2, \ldots, s_k t_k).$$

In this way, S^k is a semigroup. ■

Finitary properties have been well studied for direct product of groups. For such results classifying finite generation and finite presentation for direct products of semigroups, we refer the reader to the work of Robertson, Ruškuc & Wiegold [26], and the later work of Araújo & Ruškuc [1].

We now give a more general property of semigroups that includes the direct product, by thinking about its substructures. This is the definition of a *subdirect product*, which is ubiquitous throughout this thesis.

Definition 1.7.3. Let S and T be two semigroups. A *subdirect product* of S with T (sometimes referred to as a subdirect product of $S \times T$) is a subsemigroup U of the direct product $S \times T$, such that the projection maps defined

$$\pi_S : U \to S := (s,t) \mapsto s,$$
$$\pi_T : U \to T := (s,t) \mapsto t$$

are surjections. In this case, we will write $U \leq_{\text{sd}} S \times T$. ■

Again, we can define a subdirect product of a finite number of n semigroups as a subsemigroup of the direct product of those n semigroups, where each projection mapping onto the i-th factor is a surjection.

Examples 1.7.4. • The direct product $S \times T$ itself is a subdirect product of semigroups S and T, as every element of S appears in some first coordinate, and every element of T appears in some second coordinate. Hence the projection maps onto S and T are surjections;

• The semigroup given by $\Delta_S := \{(s,s) : s \in S\}$ is a subdirect product of a semigroup S with itself, as every element of S appears in some first and second coordinate. Δ_S is known as the *diagonal* subdirect product;

• Let F be the group with presentation

$$\left\langle x, y \mid [xy^{-1}, x^{-1}yx] = [xy^{-1}, x^{-2}yx^2] = 1 \right\rangle.$$

Then $\langle (x, y^{-1}), (y, x), (x^{-1}, x^{-1}), (y^{-1}, y) \rangle$ is a subdirect product of F with itself, as every generator of F appears in some first and second coordinate of a generating pair. $\triangle$

We will think of being a subdirect product as more of a property of a semigroup rather than a direct construction. A particular type of constructible subdirect product is a *fiber product* of semigroups, which we now define.

Definition 1.7.5. Given semigroups S, T, U and epimorphisms $\varphi : S \to U$, $\psi : T \to U$, the *fiber product of S and T with respect to φ, ψ* is the set

$$\Pi(\varphi, \psi) := \{(s, t) \in S \times T : \varphi(s) = \psi(t)\}$$

with multiplication inherited from $S \times T$. U is called the *fiber*, or *fiber quotient* of $\Pi(\varphi, \psi)$. If V is a subdirect product of $S \times T$ which is also a fiber product, we will write $V \leq_{\text{fp}} S \times T$. ■

We now qualify in the following lemma that fiber products are indeed examples of subdirect products.

Lemma 1.7.6. *For semigroups S, T, U and epimorphisms $\varphi : S \to U$, $\psi : T \to U$, $\Pi(\varphi, \psi)$ is a subdirect product of $S \times T$.*

Proof. If (s, t), $(s', t') \in \Pi(\varphi, \psi)$, then $(s, t)(s', t') = (ss', tt') \in \Pi(\varphi, \psi)$, as

$$\varphi(ss') = \varphi(s)\varphi(s') = \psi(t)\psi(t') = \psi(tt'),$$

and hence $\Pi(\varphi, \psi)$ is a subsemigroup of $S \times T$. Moreover, for any $s \in S$, as $\varphi(s) \in U$ and ψ is surjective, there exists some $t \in T$ such that $\varphi(s) = \psi(t)$, and hence $(s, t) \in \Pi(\varphi, \psi)$. Similarly, for any $t \in T$, as φ is surjective, there exists some $s \in S$ such that $\psi(t) = \varphi(s)$, and hence $(s, t) \in \Pi(\varphi, \psi)$. Thus the projection maps

$$\pi_S : \Pi(\varphi, \psi) \to S := (s, t) \mapsto s,$$
$$\pi_T : \Pi(\varphi, \psi) \to T := (s, t) \mapsto t.$$

are surjections, and $\Pi(\varphi, \psi)$ is a subdirect product of $S \times T$. □

Examples 1.7.7. • For semigroup S and T, the direct product $S \times T$ is a fiber product. We take U to be the trivial group $\{1\}$, φ to be the constant map given by

$$\varphi : S \to U := \varphi(s) = 1,$$

and similarly ψ to be the constant map given by

$$\psi : T \to U := \psi(t) = 1.$$

• For a semigroup S, the diagonal subdirect product $\Delta_S = \{(s, s) : s \in S\}$ is a fiber product. We take $U = S$, $\varphi : S \to S$ to be the identity mapping (so that $\varphi(s) = s$), and $\psi = \varphi$. Then $\varphi(s) = \psi(t) \Leftrightarrow s = t$. △

Our previous comments relating to being able to define direct and subdirect products for n semigroups still apply for fiber products. We now state an important result relating subdirect products to fiber products, known as *Fleischer's lemma*.

Lemma 1.7.8 (Fleischer's lemma, [4, Lemma 10.1])**.** *Let S, T, U be semigroups, and let $U \leq_{\text{sd}} S \times T$. For the projection maps*

$$\pi_S : U \to S := (s, t) \mapsto s,$$
$$\pi_T : U \to T := (s, t) \mapsto t,$$

denote by σ the congruence $\ker \pi_S$ on U, and denote by ρ the congruence $\ker \pi_T$ on U. Then U is a fiber product of S with T if and only if

$$\sigma \circ \rho = \rho \circ \sigma.$$

Proof. ($\Rightarrow$) If U is a fiber product of S with T, then there exist epimorphisms $\varphi : S \to V$, $\psi : T \to V$ onto a common image V, such that

$$U = \{(s, t) \in S \times T : \varphi(s) = \psi(t)\}.$$

Let $((s_1, t_1), (s_2, t_2)) \in \sigma \circ \rho$. Then there exists $(s_3, t_3) \in U$ such that

$$((s_1, t_1), (s_3, t_3)) \in \sigma, \quad ((s_3, t_3), (s_2, t_2)) \in \rho.$$

By the definitions of σ and ρ, it must be that $s_1 = s_3$, and $t_3 = t_2$. Hence $(s_1, t_2) \in U$, and thus $\varphi(s_1) = \psi(t_2)$. As $(s_1, t_1) \in U$ and $(s_2, t_2) \in U$, it follows that

$$\varphi(s_2) = \psi(t_2) = \varphi(s_1) = \psi(t_1),$$

and hence $(s_2, t_1) \in U$ also. As

$$((s_1, t_1), (s_2, t_1)) \in \rho, \quad ((s_2, t_1), (s_2, t_2)) \in \sigma,$$

then $((s_1, t_1), (s_2, t_2)) \in \rho \circ \sigma$. As $((s_1, t_1), (s_2, t_2))$ was an arbitrary element of $\sigma \circ \rho$, then we have shown that $\sigma \circ \rho \subseteq \rho \circ \sigma$. The reverse inclusion follows by a symmetric argument, and hence

$$\sigma \circ \rho = \rho \circ \sigma.$$

($\Leftarrow$) If $\sigma \circ \rho = \rho \circ \sigma$, then $\sigma \circ \rho$ is a congruence by Lemma 1.3.7. Let $\iota : U \to U/(\sigma \circ \rho)$ be the natural quotient mapping $\iota(u, v) = [(u, v)]_{\sigma\circ\rho}$. As $\sigma \subseteq \sigma \circ \rho$, there is a natural epimorphism from U/σ to $U/(\sigma \circ \rho)$ given by

$$\pi : U/\sigma \to U/(\sigma \circ \rho) := [(u, v)]_\sigma \mapsto [(u, v)]_{\sigma\circ\rho}.$$

As U is a subdirect product, then $S \cong U/\sigma$ by the first isomorphism theorem. Hence there exists an epimorphism $\varphi : S \to U/(\sigma\circ\rho)$ with $\iota = \pi_S \circ \varphi$ (recalling that our convention is to compose from left to right). A similar proof shows that there also exists an epimorphism $\psi : T \to U/(\sigma \circ \rho)$ with $\iota = \pi_T \circ \psi$. We have the following commuting diagram.

$$\begin{array}{ccc} U & \xrightarrow{\pi_T} & T \\ \downarrow{\scriptstyle \pi_S} & \overset{\iota}{\searrow} & \downarrow{\scriptstyle \psi} \\ S & \xrightarrow{\varphi} & U/(\sigma \circ \rho) \end{array}$$

We claim that

$$U = \{(u, v) \in S \times T : \varphi(u) = \psi(v)\}$$

If $(u, v) \in U$, then $\iota(u, v) = (\pi_S \circ \varphi)(u, v) = \varphi(u)$, and $\iota(u, v) = (\pi_T \circ \psi)(u, v) = \psi(v)$. Hence $\varphi(u) = \psi(v)$.

Conversely, if $(u, v) \in S \times T$ is such that $\varphi(u) = \psi(v)$, then as U is a subdirect product of S with T, there exist $s \in S$ and $t \in T$ such that $(u, t) \in U$ and $(s, v) \in U$. Now

$$\iota(u, t) = (\pi_S \circ \varphi)(u, v) = \varphi(u) = \psi(v) = (\pi_T \circ \psi)(s, v) = \iota(s, v),$$

and hence $((u, t), (s, v)) \in \sigma \circ \rho$. Hence there must exist $(x, y) \in U$ such that $((u, t), (x, y)) \in \sigma$ and $((x, y), (s, v)) \in \rho$. This implies however that $x = u$ and $y = v$, and hence $(u, v) \in U$. Thus we have shown that

$$U = \{(u, v) \in S \times T : \varphi(u) = \psi(v)\},$$

and U is a fiber product. □

In varieties of algebras for which the composition of congruences is a commutative operation (such as in groups, where congruences are equivalent to normal subgroups), then fiber products and subdirect products are one and the same as a consequence of Fleischer's lemma. Such varieties are called *congruence permutable*. The varieties of semigroups and monoids are not congruence permutable however, as semigroup and monoid congruences do not commute under composition in general.

For the rest of this section, we outline some of the previous results (sans proof) in the theory of subdirect products for groups that we motivate this thesis with, as discussed in the introduction. These all notably involve subdirect products of free groups.

The first result we highlight is a corollary of Bridson & Miller [3], which finds uncountably many non-isomorphic subdirect products of two free groups.

Theorem 1.7.9 ([3, Corollary B])**.** *Let F_1 and F_2 be non-abelian free groups. Then there are uncountably many subdirect products G of $F_1 \times F_2$, with non-isomorphic first cohomology groups (with associated G-module $\mathbb{Z}$).* □

As a consequence of this result, as two isomorphic groups have isomorphic first cohomology groups, then $F_1 \times F_2$ contains uncountably many non-

isomorphic subdirect products. This gives a generalisation of the work of Baumslag & Roseblade [2] on direct products of free groups of rank two, which we highlight below.

Theorem 1.7.10 ([2, Theorem 1]). *There are uncountably many non-isomorphic subgroups of the direct product of two free groups of rank two.* □

In particular, the authors prove this using subdirect products of two free groups of rank two. These results in part motivate our study of the number of non-isomorphic subdirect products of two free monogenic semigroups in Chapter 2.

Turning now to examples involving finitary properties, we first note that many finitary properties of groups are preserved under taking the direct product. That is, the statement

$$G \times H \text{ has property } P \text{ if and only if } G \text{ and } H \text{ have property } P$$

is satisfied for P being the properties of being finitely generated, finitely presented, residually finite, amongst others. It is also satisfied for many other interesting non-finitary properties, such as nilpotency. Further generalisations of such statements for direct products of semigroups are considered by Robertson, Ruškuc & Wiegold [26, Theorem 2.1, Theorem 8.3] in the cases of finite generation and finite presentation respectively, and by Gray & Ruškuc [12, Theorem 2] in the case of residual finiteness.

This is not necessarily the case for subdirect products however. In the work of Bridson & Miller, an example is given showing that finite presentation of direct product factors cannot guarantee finite generation of a subdirect product of those factors. We present this example for motivation, without proof, referring the reader to [3, Example 3] for further detail.

Example 1.7.11 ([3, Example 3]). Let $A = \langle a_1, a_2 : \emptyset \rangle$ and $B = \langle b_1, b_2 : \emptyset \rangle$ be two free groups of rank 2, and let

$$Q = \langle c_1, c_2 : q_1(c), q_2(c), \dots \rangle$$

be any two-generated group which is not finitely presented. Let $\varphi : A \to Q$ and $\psi : B \to Q$ be the unique epimorphisms extending the maps defined by $\varphi(a_i) = c_i$ and $\psi(b_i) = c_i$. The fiber product $\Pi(\varphi, \psi)$ of $A \times B$ is not finitely generated. $\triangle$

Further, Grunewald [13] has classified when subdirect products of free groups are finitely presented. This was reproved later by Baumslag & Roseblade [2, Theorem 2], with the following generalisation.

Theorem 1.7.12 ([2, Theorem 2])**.** *Every finitely presented subgroup of a direct product of two free groups is a finite extension of a direct product of two free groups of finite rank.* $\square$

Not every finitely generated subdirect product of two free groups has this property, and hence there exist finitely generated subdirect products which are not finitely presented.

On a similar note, a final example we mention is due to Mikhaĭlova [21]. In this work a set of defining relations is given for a group which is a finitely generated subdirect product of two free groups, but has undecidable word problem. This is yet another example involving subdirect products of free groups, that motivate our study in this thesis of subdirect products of free semigroups and monoids.

Chapter 2

Subdirect products involving the free monogenic semigroup

As we saw in Chapter 1, Baumslag & Roseblade showed that the direct product of two free groups can have uncountably many subgroups up to isomorphism (Theorem 1.7.10), which was then extended to subdirect products of non-abelian free groups by Bridson & Miller (Theorem 1.7.9). This illustrates some of the sub-structural complexity that the direct product can create (and in particular, the fiber product construction for groups).

Notably, their result used the direct product of two free groups of rank at least 2, and we recall from Section 1.7 that some examples of those subgroups can be non-finitely generated, finitely generated whilst not being finitely presented, and being finitely generated whilst having the membership problem being undecidable. In particular, we indicated that these examples include subdirect products, which perhaps gives us motivation to believe that subdirect products involving free semigroups can provide further instances of interesting and erratic structural behaviour.

By way of contrast, however, subgroups and subdirect products of the direct product of two free groups of rank 1 (that is, the free cyclic group $\mathbb{Z}$ under addition) do not exhibit this same behaviour. The group $\mathbb{Z} \times \mathbb{Z}$ is the unique (up to isomorphism) free abelian group of rank two, and hence every non-

trivial subgroup of $\mathbb{Z} \times \mathbb{Z}$ is free abelian, and of rank less than or equal to two (this is illustrated and proved in [18]). Hence its non-trivial subgroups are isomorphic to either $\mathbb{Z}$, or $\mathbb{Z} \times \mathbb{Z}$.

As a corollary, every subgroup of $\mathbb{Z} \times \mathbb{Z}$ is therefore finitely presented and has decidable membership problem, and of course there are only countably many of them up to isomorphism. That every subgroup of $\mathbb{Z}$ is isomorphic to either the trivial subgroup or $\mathbb{Z}$ itself is a major contributing factor for these resulting "well-behaved" properties.

By way of comparison, the corresponding free monogenic object in the variety of semigroups is isomorphic to the natural numbers $\mathbb{N}$ under the operation of addition. Similarly to $\mathbb{Z}$, its subsemigroups are well understood due to the following result of Sit & Siu [27].

Theorem 2.0.1 (Sit & Siu, [27])**.** *Every subsemigroup M of $\mathbb{N}$ is of the form $M = X \cup Y$, where $X \subseteq \{1, \ldots, N\}$, and $Y = \{m \in \mathbb{N} : m \geq N, d \mid m\}$ for some $d, N \in \mathbb{N}$.* □

As with $\mathbb{Z}$ and $\mathbb{Z} \times \mathbb{Z}$, the subsemigroups (and hence subdirect products) of $\mathbb{N}$ are all finitely presented and there are only countably many of them up to isomorphism. It is for these reasons and those outlined above that we investigate the subdirect products and subsemigroups of $\mathbb{N} \times \mathbb{N}$ in this chapter.

Reflecting on the work of Baumslag & Roseblade [2], and Bridson & Miller [3] on free groups of rank at least two, our main aim for this chapter will be to prove that $\mathbb{N} \times \mathbb{N}$ in fact has uncountably many subsemigroups and subdirect products up to isomorphism in Section 2.1, and derive related corollaries about semigroups with elements of infinite order such as $\mathbb{N}^k$ in Section 2.3. We will go on to investigate subsemigroups and subdirect products of $\mathbb{N}$ together with a finite semigroup in Sections 2.2 and 2.4 respectively; classifying for which finite semigroups S there are uncountably many subsemigroups and subdirect products of $\mathbb{N} \times S$.

We note that this chapter is largely based on the paper [7], cowritten by the

author.

2.1 Subsemigroups of $\mathbb{N} \times \mathbb{N}$

In this section, we will see that the direct product $\mathbb{N} \times \mathbb{N}$ contains uncountably many subsemigroups up to isomorphism, and we will further derive a corollary about the number of subsemigroups up to isomorphism of the direct product of two infinite semigroups. Later, in Section 2.3, we will draw parallel results on the number of non-isomorphic subdirect powers of $\mathbb{N}$.

We begin by adopting notation for the purposes of this section.

Notation 2.1.1. Given a non-empty subset $M \subseteq \mathbb{N}$, we will define X_M to be the subset $\{1\} \times M$ of $\mathbb{N} \times \mathbb{N}$, and S_M to be the subsemigroup of $\mathbb{N} \times \mathbb{N}$ generated by the set X_M. ■

We comment that $(m, n) \in S_M$ if and only if n can be written as the sum of exactly m (not necessarily distinct) numbers from M.

Examples 2.1.2. (a). Let $M = \{1\}$. Then X_M consists of the single pair $(1, 1)$, and hence $S_M = \langle (1, 1) \rangle = \{(n, n) : n \in \mathbb{N}\}$. In this case, S_M is isomorphic to $\mathbb{N}$, via the natural mapping $(m, m) \mapsto m$.

(b). Let $N = \{1, 2\}$. Then $X_N = \{(1, 1), (1, 2)\}$, and

$$S_N = \{(m + n, m + 2n) : m, n \in \mathbb{N}^0\} \setminus \{(0, 0)\}. \qquad \triangle$$

In fact, we will see that S_N from Examples 2.1.2 (b) is isomorphic to the free commutative semigroup of rank 2, as the following result documents.

Lemma 2.1.3. *S_M is a free commutative semigroup if and only if $|M| \leq 2$.*

Proof. If $|M| = 1$, then $M = \{m\}$ for some $m \in \mathbb{N}$. Hence

$$S_M = \langle (1, m) \rangle = \{(n, nm) : n \in \mathbb{N}\}.$$

Then the mapping $\varphi : S_M \to \mathbb{N} := (n, nm) \mapsto n$ is clearly bijective, and

$$\varphi((n, nm)+(p, pm)) = \varphi((n+p, (n+p)m)) = n+p = \varphi((n, nm))+\varphi((p, pm)),$$

and hence is an isomorphism from S_M to the free commutative monogenic semigroup.

If $|M| = 2$, let $M = \{m_1, m_2\}$. It suffices to show that any given relation in S_M is trivial. As a given relation in S_M can be viewed as an equivalence of two non-zero sums of elements from X_M, then it is of the form

$$p(1, m_1) + q(1, m_2) = r(1, m_1) + s(1, m_2) \tag{2.1}$$

for some $p, q, r, s \in \mathbb{N}^0$ (not all zero). Simplifying and equating components, we obtain the relations

$$\begin{aligned} p + q &= r + s, \\ pm_1 + qm_2 &= rm_1 + sm_2, \end{aligned}$$

which are equivalent to the integer equations

$$\begin{aligned} p - r &= s - q, \\ (p - r)m_1 &= (s - q)m_2. \end{aligned}$$

As $m_1 \neq m_2$ by assumption, to avoid contradiction it must be that $p - r = s - q = 0$, and hence $p = r, q = s$ and the relation given in (2.1) is trivial. Hence S_M is a free commutative semigroup, with rank 2.

If $|M| \geq 3$, let $m_1, m_2, m_3 \in M$ be any three distinct naturals. Then

$$\begin{aligned} &m_2(1, m_1) + m_3(1, m_2) + m_1(1, m_3) = \\ &m_3(1, m_1) + m_1(1, m_2) + m_2(1, m_3) \end{aligned} \tag{2.2}$$

is a non-trivial relation between the generators X_M, and hence S_M cannot be free commutative. □

As a consequence of the preceding result, there are only two non-isomorphic semigroups in the family $\{S_M : M \subset \mathbb{N}, |M| \leq 2\}$, and hence for the rest

of the chapter we will only consider the semigroups S_M where $|M| \geq 3$. We will also utilise the relation (2.2) later when discussing 3-element subsets. We will largely consider their possible isomorphisms by the images of their indecomposable elements, which we now describe.

Lemma 2.1.4. *The indecomposable elements of S_M are exactly the elements of X_M.*

Proof. If an element $(p, q) \in S_M$ has $p > 1$, then (p, q) is necessarily a sum of p elements of X_M by construction of S_M, and is decomposable. Hence the indecomposable elements must be contained within X_M, as these are the only elements of S_M with first component equal to 1.

Conversely, as the elements of X_M have first coordinate equal to 1, then they are indecomposable in $\mathbb{N} \times \mathbb{N}$, and hence indecomposable in S_M. □

For two 3-element subsets $M = \{m_1, m_2, m_3\}$, $N = \{n_1, n_2, n_3\}$ of $\mathbb{N}$, any isomorphism from S_M to S_N must map X_M to X_N by Lemma 1.6.11. Hence without loss of generality, we can fix the labelling of M and N such that a given isomorphism $\varphi : S_M \to S_N$ satisfies $\varphi(1, m_i) = (1, n_i)$ for $i = 1, 2, 3$.

We now use this fact in the next result to classify when two 3-element subsets $M, N \subseteq \mathbb{N}$ give rise to isomorphic semigroups S_M and S_N.

Lemma 2.1.5. *Let $M = \{m_1, m_2, m_3\}$, $N = \{n_1, n_2, n_3\}$ be two 3-element subsets of $\mathbb{N}$. Then there is an isomorphism $\varphi : S_M \to S_N$ satisfying $\varphi(1, m_i) = (1, n_i)$ $(i = 1, 2, 3)$, if and only if*

$$n_2(m_3 - m_1) = n_1(m_3 - m_2) + n_3(m_2 - m_1). \tag{2.3}$$

Proof. ($\Rightarrow$) Suppose that $S_M \cong S_N$ via the isomorphism φ. As noted in Lemma 2.1.3 equation (2.2), the relation

$$m_2(1, m_1) + m_3(1, m_2) + m_1(1, m_3) = m_3(1, m_1) + m_1(1, m_2) + m_2(1, m_3)$$

holds in S_M. Applying φ to the left hand side of this relation, we see that

$$\begin{aligned}&\varphi(m_2(1,m_1)+m_3(1,m_2)+m_1(1,m_3))\\&=m_2\varphi(1,m_1)+m_3\varphi(1,m_2)+m_1\varphi(1,m_3)\\&=m_2(1,n_1)+m_3(1,n_2)+m_1(1,n_3).\end{aligned}\tag{2.4}$$

Similarly, applying φ to the right hand side of (2.2) gives

$$\begin{aligned}&\varphi(m_3(1,m_1)+m_1(1,m_2)+m_2(1,m_3))\\&=m_3\varphi(1,m_1)+m_1\varphi(1,m_2)+m_2\varphi(1,m_3)\\&=m_3(1,n_1)+m_1(1,n_2)+m_2(1,n_3).\end{aligned}\tag{2.5}$$

Simplifying (2.4) and (2.5) and equating second components gives the equation

$$m_2n_1+m_3n_2+m_1n_3=m_3n_1+m_1n_2+m_2n_3,$$

from which (2.3) is obtained after factorisation of n_1, n_2 and n_3.

($\Leftarrow$) Assume that (2.3) holds. As any element $x \in S_M$ can be written as a linear combination of elements from X_M, let

$$x=\alpha_1(1,m_1)+\alpha_2(1,m_2)+\alpha_3(1,m_3)\in S_M$$

be such a combination, for some $\alpha_1, \alpha_2, \alpha_3 \in \mathbb{N}^0$, not all zero. Define the mapping $\varphi : S_M \to S_N$ on such a linear combination by

$$\varphi(x):=\alpha_1(1,n_1)+\alpha_2(1,n_2)+\alpha_3(1,n_3).$$

As the factors $\alpha_1, \alpha_2, \alpha_3$ are not necessarily uniquely determined by x as seen in (2.2), we first show that φ is well defined. To this end, suppose $\alpha_i, \beta_i \in \mathbb{N}_0$ (not all zero, $i = 1, 2, 3$) are such that

$$\alpha_1(1,m_1)+\alpha_2(1,m_2)+\alpha_3(1,m_3)=\beta_1(1,m_1)+\beta_2(1,m_2)+\beta_3(1,m_3).$$

Let $\gamma_i = \alpha_i - \beta_i$. Noting that (2.3) is equivalent to the relation

$$n_1(m_2 - m_3) + n_2(m_3 - m_1) + n_3(m_1 - m_2) = 0, \tag{2.6}$$

then

$$\begin{aligned}
& \textstyle\sum_{i=1}^3 \alpha_i(1, m_i) = \sum_{i=1}^3 \beta_i(1, m_i) \\
\Rightarrow\ & \textstyle\sum_{i=1}^3 \alpha_i = \sum_{i=1}^3 \beta_i \text{ and } \sum_{i=1}^3 \alpha_i m_i = \sum_{i=1}^3 \beta_i m_i \\
\Rightarrow\ & \textstyle\sum_{i=1}^3 \gamma_i = 0 \text{ and } \sum_{i=1}^3 \gamma_i m_i = 0 \\
\Rightarrow\ & \gamma_2 m_2 = -\gamma_1 m_2 - \gamma_3 m_2,\ \gamma_2 m_2 = -\gamma_1 m_1 - \gamma_3 m_3, \\
& \gamma_1 m_1 = -\gamma_2 m_1 - \gamma_3 m_1 \text{ and } \gamma_1 m_1 = -\gamma_2 m_2 - \gamma_3 m_3 \\
\Rightarrow\ & \textstyle\sum_{i=1}^3 \gamma_i = 0,\ \gamma_1(m_1 - m_2) = \gamma_3(m_2 - m_3) \\
& \text{and } \ \gamma_2(m_1 - m_2) = \gamma_3(m_3 - m_1) \\
\Rightarrow\ & \textstyle\sum_{i=1}^3 \gamma_i = 0 \text{ and } \sum_{i=1}^3 \gamma_i n_i(m_1 - m_2) = \gamma_3(0) = 0 \text{ (by (2.6))} \\
\Rightarrow\ & \textstyle\sum_{i=1}^3 \gamma_i = 0 \text{ and } \sum_{i=1}^3 \gamma_i n_i = 0 \text{ (as } m_1 - m_2 \neq 0) \\
\Rightarrow\ & \textstyle\sum_{i=1}^3 \alpha_i = \sum_{i=1}^3 \beta_i \text{ and } \sum_{i=1}^3 \alpha_i n_i = \sum_{i=1}^3 \beta_i n_i \\
\Rightarrow\ & \textstyle\sum_{i=1}^3 \alpha_i(1, n_i) = \sum_{i=1}^3 \beta_i(1, n_i).
\end{aligned}$$

It follows that φ is well defined, as now

$$\varphi\big(\sum_{i=1}^3 \alpha_i(1, m_i)\big) = \sum_{i=1}^3 \alpha_i(1, n_i) = \sum_{i=1}^3 \beta_i(1, n_i) = \varphi\big(\sum_{i=1}^3 \beta_i(1, m_i)\big).$$

Moreover, noting that (2.6) is equivalent to

$$m_1(n_2 - n_3) + m_2(n_3 - n_1) + m_3(n_1 - n_2) = 0, \tag{2.7}$$

by rearrangement, then swapping the roles of m and n, and the roles (2.6) and (2.7) in the above series of implications also shows that φ is injective. That it is a homomorphism and surjective follows directly from definition, and thus φ is an isomorphism between S_M and S_N. □

Examples 2.1.6. (*a*). Let $M = \{1, 2, 3\}, N = \{2, 4, 6\}$. We expect S_M and S_N to be isomorphic, as $N = 2M$. Indeed, as $4(3-1) = 2(3-2) + 6(2-1)$,

then setting $m_i = i$ and $n_i = 2i$ for $i = 1, 2, 3$ gives

$$n_2(m_3 - m_1) = n_1(m_3 - m_2) + n_3(m_2 - m_1).$$

Hence by Lemma 2.1.5, S_M and S_N are isomorphic, with isomorphism φ satisfying $\varphi(1, i) = (1, 2i)$ for $i = 1, 2, 3$.

(b). Let $m_i = i$ for $i = 1, 2, 3$, $n_1 = 2, n_2 = 1, n_3 = 3$, and let $M = \{m_1, m_2, m_3\}, N = \{n_1, n_2, n_3\}$. We will show that S_M and S_N are isomorphic, but the attempt at constructing a homomorphism $\varphi : S_M \to S_N$ satisfying

$$\varphi(1, 1) = (1, 2),\ \varphi(1, 2) = (1, 1),\ \varphi(1, 3) = (1, 3)$$

does not give an isomorphism. Indeed, note that φ is not even well-defined, as

$$2(1, 1) + 3(1, 2) + (1, 3) = (6, 11) = 3(1, 1) + (1, 2) + 2(1, 3),$$

but

$$\varphi(2(1, 1) + 3(1, 2) + (1, 3)) = 2(1, 2) + 3(1, 1) + (1, 3) = (6, 10),$$

and

$$\varphi(3(1, 1) + (1, 2) + 2(1, 3)) = 3(1, 2) + (1, 1) + 2(1, 3) = (6, 13).$$

This is verified by Lemma 2.1.5, as $1(3 - 1) \neq 2(3 - 2) + 3(2 - 1)$. However, S_M and S_N are clearly isomorphic, as $M = N$. An example of a non-trivial isomorphism between S_M and S_N is $\psi : S_M \to S_N$ satisfying

$$\psi(1, 1) = (1, 3),\ \psi(1, 2) = (1, 2),\ \psi(1, 3) = (1, 1),$$

as $2(3 - 1) = 3(3 - 2) + 1(2 - 1)$.

(c). Let $M = \{1, 2, 3\}, N = \{1, 2, 100\}$. Suppose there were a labelling of M and N such that (2.3) holds. If $n_1 = 100$, then as (2.3) is equivalent to the condition

$$n_1(m_2 - m_3) = n_2(m_1 - m_3) + n_3(m_2 - m_1),$$

then it follows from the triangle inequality that

$$100|m_2 - m_3| \leq n_2|m_1 - m_3| + n_3|m_2 - m_1|.$$

But the differences $|m_i - m_j|$ are at most 2, and hence it can be quickly verified that right hand side of this inequality is at most 5, which is a contradiction. Hence $n_1 \neq 100$. The same contradiction is obtained using (2.3) supposing $n_2 = 100$, and a similar contradiction can be obtained from rearranging (2.3) supposing that $n_3 = 100$. Hence there is no labelling of M and N such that (2.3) holds, and hence S_M cannot be isomorphic to S_N. △

Lemma 2.1.5 and the relation (2.3) will be our main tools in obtaining uncountably many pairwise non-isomorphic semigroups S_M, and hence motivate the following definition.

Definition 2.1.7. A subset $M \subseteq \mathbb{N}$ will be called *3-separating* if $|M| \geq 3$, and additionally, for any two triples (m_1, m_2, m_3) and (n_1, n_2, n_3) of distinct elements from M, the following holds:

$$\begin{aligned} &n_2(m_3 - m_1) = n_1(m_3 - m_2) + n_3(m_2 - m_1) \\ \Leftrightarrow\ &(m_1, m_2, m_3) = (n_1, n_2, n_3). \end{aligned} \tag{C1}$$

■

Examples 2.1.8. (a). The set $M = \{1, 2, 3\}$ is not 3-separating, as

$$2(3 - 1) = 3(3 - 2) + 1(2 - 1),$$

but $(1, 2, 3) \neq (3, 2, 1)$.

(b). For the set $\{2, 3, 5\}$, the possible triples of distinct elements are

$$(2, 3, 5), (2, 5, 3), (3, 2, 5), (3, 5, 2), (5, 2, 3), (5, 3, 2).$$

A direct calculation of all possibilities (which we omit for brevity) verifies that the only pairs of triples (m_1, m_2, m_3) and (n_1, n_2, n_3) satisfying condition (2.3) are precisely those with $m_i = n_i$, for $i = 1, 2, 3$, and hence $\{2, 3, 5\}$ is a

3-separating set. $\triangle$

We now focus on extending such examples of finite 3-separating sets to infinite 3-separating sets. This will aid us in constructing infinite semigroups S_M for which any pair of 3-generated subsemigroups are non-isomorphic, as a consequence of Lemma 2.1.5. Our approach will be inductive, and we begin by deriving a condition on pairs of elements from 3-separating sets which guarantees that the induction holds.

Lemma 2.1.9. *Let M be a 3-separating set. Then for any two pairs (m_1, m_2), (n_1, n_2) of distinct elements from M, the following condition holds:*

$$m_1 - m_2 = n_1 - n_2 \Leftrightarrow (m_1, m_2) = (n_1, n_2). \tag{C2}$$

Proof. We prove the contrapositive. Suppose that there exist two pairs (m_1, m_2), (n_1, n_2) of distinct elements from M with $(m_1, m_2) \neq (n_1, n_2)$, but

$$m_1 - m_2 = n_1 - n_2. \tag{2.8}$$

It follows that $m_1 \neq n_1$, as otherwise (2.8) gives that $m_2 = n_2$. Similarly, $m_2 \neq n_2$. This leaves the following three cases: either $m_1 = n_2$; $m_2 = n_1$; or $\{m_1, m_2, n_1, n_2\}$ is a set of four distinct naturals. In each case, we show that M is not 3-separating:

Case 1: $m_1 = n_2$. In this instance $m_2 \neq n_1$, for otherwise $m_1 - m_2 = n_2 - n_1 = n_1 - n_2$, and hence $n_1 = n_2$ which is a contradiction. Thus the two triples (n_1, m_1, m_2) and (m_2, m_1, n_1) are non-equal and consist of distinct elements from M. We will further show that they satisfy the left hand side of condition (C1).

Letting $d = m_1 - n_1$, then it also follows from (2.8) that $d = m_2 - m_1$, and hence the triples (n_1, m_1, m_2) and (m_2, m_1, n_1) may be re-expressed as $(n_1, n_1 + d, n_1 + 2d)$ and $(n_1 + 2d, n_1 + d, n_1)$ respectively. Thus considering (C1), we have

$$m_2(m_2 - m_1) + n_1(m_1 - n_1)$$

$$
\begin{aligned}
&= (n_1 + 2d)(n_1 + 2d - n_1 - d) + n_1(n_1 + d - n_1) \\
&= (n_1 + 2d)d + dn_1 \\
&= 2d^2 + 2dn_1 \\
&= (n_1 + d)(n_1 + 2d - n_1) \\
&= m_1(m_2 - n_1),
\end{aligned}
$$

and hence the two non-equal triples (n_1, m_1, m_2) and (m_2, m_1, n_1) violate (C1), and M is not 3-separating.

Case 2: $m_2 = n_1$. This case can be handled identically to Case 1, reversing the roles of the indices 1 and 2.

Case 3: $\{m_1, m_2, n_1, n_2\}$ *is a set of four distinct naturals.* We claim that the two triples (n_2, n_1, m_2) and (m_1, m_2, n_1) of distinct elements of M satisfy the left hand side of (C1).

Let $d = m_1 - m_2 = n_1 - n_2$. Then the two triples (n_2, n_1, m_2) and (m_1, m_2, n_1) can be re-expressed as $(n_2, n_2 + d, m_1 - d)$ and $(m_1, m_1 - d, n_2 + d)$. Thus considering (C1), we have

$$
\begin{aligned}
&(m_1 - d)(m_1 - d - n_2) \\
&= m_1^2 - 2dm_1 + d^2 - n_2 m_1 + dn_2 \\
&= m_1(m_1 - d - (n_2 + d)) + (n_2 + d)(n_2 + d - n_2)
\end{aligned}
$$

and hence the two non-equal triples (n_2, n_1, m_2) and (m_1, m_2, n_1) violate (C1), and M is not 3-separating. □

We remark that any subset N (with $|N| \geq 3$) of a 3-separating set M is 3-separating. Hence constructing a countably infinite 3-separating set will yield uncountably many 3-separating subsets. To start this construction, we next show that finite 3-separating sets can be added to, whilst keeping the 3-separating property.

Lemma 2.1.10. *If M is a 3-separating finite set, then there exists $x \in \mathbb{N} \setminus M$ such that $M \cup \{x\}$ is also 3-separating.*

Proof. If $M \cup \{x\}$ is not 3-separating for some $x \in \mathbb{N} \setminus M$, we will show that there are only finitely many such possibilities for x. This will complete the proof of the lemma, as we are still left with infinitely many choices of x to pick from $\mathbb{N} \setminus M$ so that $M \cup \{x\}$ is 3-separating.

As $M \cup \{x\}$ is assumed to not be a 3-separating set, there exists a pair of triples $(m_1, m_2, m_3), (n_1, n_2, n_3)$ of distinct elements from $M \cup \{x\}$ with $(m_1, m_2, m_3) \neq (n_1, n_2, n_3)$, but

$$n_2(m_3 - m_1) = n_1(m_3 - m_2) + n_3(m_2 - m_1). \tag{2.9}$$

At least one of the m_i, n_i is equal to x as M is 3-separating. At most one of the m_i can be equal to x, and at most one of the n_j can be equal to x, giving us the following three possible cases:

Case 1: Exactly one of $m_1, m_2, m_3, n_1, n_2, n_3$ is equal to x. In this instance, (2.9) can be regarded as a linear equation in x with non-zero coefficient. Hence for a given choice of the m_i, n_i which are not equal to x, there is at most one such x such that (2.9) holds. As there are only finitely many choices for the five of $m_1, m_2, m_3, n_1, n_2, n_3$ which are not equal to x, there are only finitely many such x violating (2.9).

Case 2: $m_i = n_j = x$ for some distinct $i, j \in \{1, 2, 3\}$. In this instance, (2.9) can be rearranged into a quadratic equation of x, in particular with non-zero quadratic coefficient. There will be at most two solutions for x for any choice of the four of $m_1, m_2, m_3, n_1, n_2, n_3$ which are not equal to x, and hence again there are only finitely many such x violating (2.9).

Case 3: $m_i = n_i = x$ for some $i = 1, 2, 3$. As in Case 1, (2.9) can be rearranged into a linear equation in x, with the coefficient of x being of the form $m_j - m_k + n_k - n_j$ for $j, k \in \{1, 2, 3\} \setminus \{i\}$. If this coefficient were zero, then $m_j - m_k = n_j - n_k$, and hence as m_j, m_k, n_j, n_k are elements of M which is 3-separating, then $(m_j, m_k) = (n_j, n_k)$ by Lemma 2.1.9. But then $m_l = n_l$ for each $l = 1, 2, 3$ which contradicts the choice of (m_1, m_2, m_3) and (n_1, n_2, n_3).

Hence again in this instance, (2.9) can be regarded as a linear equation in x with non-zero coefficient, there will be at most one solution for x for any choice of the four of $m_1, m_2, m_3, n_1, n_2, n_3$ which are not equal to x, and thus only finitely many x violate (2.9). □

We now construct an infinite 3-separating set in the following corollary.

Corollary 2.1.11. *There exists an infinite* 3*-separating set* M_∞*.*

Proof. Let M_1 be any finite 3-separating set (such as in Examples 2.1.8 (b)). By Lemma 2.1.10, there exists an $x_1 \in \mathbb{N} \setminus M_1$ such that $M_2 = M_1 \cup \{x_1\}$ is 3-separating.

Continuing iteratively, we obtain an infinite strictly ascending chain

$$M_1 \subset M_2 \subset M_3 \subset \dots$$

of finite 3-separating sets. Letting $M_\infty = \bigcup_{i \in \mathbb{N}} M_i$, we claim that M_∞ is 3-separating. Suppose for a contradiction that M_∞ is not 3-separating. Then there would exist two non-equal triples (m_1, m_2, m_3) and (n_1, n_2, n_3) of distinct elements from M_∞ violating (C1). However, any finite subset of M_∞ is contained within some M_i, and hence $\{m_1, m_2, m_3, n_1, n_2, n_3\}$ is contained within some M_i which is 3-separating, contradicting the choice of triples. Hence M_∞ is an infinite 3-separating. □

The existence of M_∞ from Corollary 2.1.11 is enough to prove the main theorem of this section, which we now present.

Theorem 2.1.12. *There are uncountably many pairwise non-isomorphic subsemigroups of* $\mathbb{N} \times \mathbb{N}$*.*

Proof. Let M_∞ be an infinite 3-separating set, such as in Corollary 2.1.11. We claim that any two semigroups in the collection

$$\mathcal{C} = \{S_M : M \subseteq M_\infty, |M| \geq 3\} \tag{2.10}$$

are non-isomorphic. Suppose for a contradiction that S_M and S_N were isomorphic, with $M, N \subseteq M_\infty$, $|M|, |N| \geq 3$, but $M \neq N$, via isomorphism $\varphi : S_M \to S_N$.

Without loss of generality, we can assume that $M \setminus N$ is non-empty, and hence we can choose some $m_1 \in M \setminus N$. Let m_2, m_3 be any other distinct elements from M.

The elements $(1, m_i)$ for $i = 1, 2, 3$ belong to X_M, and are thus indecomposable in S_M by Lemma 2.1.4. As the images of $(1, m_i)$ must be indecomposable in S_N by Lemma 1.6.11, and the indecomposables of S_N are the set X_N, then each $(1, m_i)$ must be mapped to $(1, n_i)$ for some $n_i \in N$, where $i = 1, 2, 3$.

Hence the subsemigroups

$$\langle (1, m_1), (1, m_2), (1, m_3) \rangle \leq S_M,$$
$$\langle (1, n_1), (1, n_2), (1, n_3) \rangle \leq S_N$$

are isomorphic via the restriction of φ to $X_{\{m_1, m_2, m_3\}}$. But by Lemma 2.1.5,

$$n_2(m_3 - m_1) = n_1(m_3 - m_2) + n_3(m_2 - m_1).$$

As M_∞ is 3-separating, it must be that $(m_1, m_2, m_3) = (n_1, n_2, n_3)$ by (C1). However this is a contradiction, as $m_1 = n_1 \in N$, but m_1 was chosen to be in $M \setminus N$. Hence $S_M \ncong S_N$ if $M \neq N$. As there are uncountably many subsets of M_∞, the result of the theorem follows. □

As $\mathbb{N}$ is an infinite monogenic semigroup, we can obtain the following corollary about the number of subsemigroups of the product of two infinite semigroups.

Corollary 2.1.13. *If S and T are infinite semigroups, each containing an element of infinite order, then the direct product of S and T contains uncountably many pairwise non-isomorphic subsemigroups.*

Proof. Let $s \in S$, $t \in T$ be elements of infinite order. Then as $\langle x \rangle \cong \langle y \rangle \cong \mathbb{N}$, it follows that $S \times T$ contains the subsemigroup $\langle x \rangle \times \langle y \rangle$, which is isomorphic

to $\mathbb{N} \times \mathbb{N}$. As $\mathbb{N} \times \mathbb{N}$ contains uncountably many pairwise non-isomorphic subsemigroups by Theorem 2.1.12, so too does $S \times T$. □

As the final result of this section, we note that Theorem 2.1.12 can be generalised to finite direct powers $\mathbb{N}^k$.

Theorem 2.1.14. *For $k \geq 2$, the direct power $\mathbb{N}^k$ contains uncountably many pairwise non-isomorphic subsemigroups.*

Proof. Let M_∞ be an infinite 3-separating set, such as in Corollary 2.1.11. Then for any $M \subseteq M_\infty$ with $|M| \geq 3$, define the set $Y_M = \{(1, \ldots, 1, m) : m \in M\}$, and the subsemigroup $T_M = \langle Y_M \rangle \leq \mathbb{N}^k$. Then the map

$$\varphi : T_M \to S_M := (n, \ldots, n, p) \mapsto (n, p)$$

is verifiably an isomorphism between T_M and S_M, and hence the result follows from Corollary 2.1.11. □

2.2 Subsemigroups of direct products of $\mathbb{N}$ with a finite semigroup

In the previous section, we saw that although $\mathbb{N}$ has only countably many subsemigroups up to isomorphism, finite direct powers of $\mathbb{N}$ (in particular, $\mathbb{N} \times \mathbb{N}$) have uncountably many. It is also the case for two finite semigroups S, T that the direct product $S \times T$ has only finitely many subsemigroups up to isomorphism (as $|S \times T| = |S| \times |T|$). It is perhaps natural to then ask the same question for direct products involving $\mathbb{N}$ that are somehow "inbetween" $\mathbb{N}$ and $\mathbb{N} \times \mathbb{N}$.

In this section, we aim to answer that question directly, characterising for which finite semigroups S does $\mathbb{N} \times S$ have countably many subsemigroups up to isomorphism. Later in Section 2.4, we will give an accompanying result on the number of non-isomorphic subdirect products.

We begin this section by dealing with the case where S is a finite group, in

the following lemma.

Lemma 2.2.1. *Let G be a finite group. Then every subsemigroup of $\mathbb{N} \times G$ is finitely generated, and hence $\mathbb{N} \times G$ has countably many subsemigroups.*

Proof. We will provide a generating set for any such $T \leq \mathbb{N} \times G$. Let $m \in \mathbb{N}$ be fixed so that $(m, 1_G) \in T$, which exists as G is a finite group, and any $(n, g) \in T$ is such that the second coordinate of $(n, g)^{|g|}$ is 1_G. For all $n \in \mathbb{N}$, we define the set

$$G_n := \{g \in G : (n, g) \in T\}.$$

Then given any $n \in \mathbb{N}$ and $g \in G_n$, it follows that

$$(n + m, g) = (n, g)(m, 1_G) \in T,$$

and thus $g \in G_{n+m}$. Hence as $G_n \subseteq G_{n+m}$ for all $n \in \mathbb{N}$, we obtain an ascending chain

$$G_n \subseteq G_{n+m} \subseteq G_{n+2m} \subseteq \dots$$

which must eventually stabilise (meaning there exists some $i \in \mathbb{N}$ with $G_{n+im} = G_{n+jm}$ for all $j \geq i$), as G is finite. Hence the sequence $(G_i)_{i \in \mathbb{N}}$ is such that there exists $j_0 \in \mathbb{N}$ with $G_j = G_{j+m}$ for all $j \geq j_0$. Fix such a j_0.

We will show that T is generated by the set

$$X = \bigcup_{1 \leq n < j_0 + m} (\{n\} \times G_n),$$

which is finite.

Firstly, as $X \subseteq T$ by construction, then the semigroup generated by X is contained in T (that is, $\langle X \rangle \subseteq T$). It remains to show that every element of T can be written as a product of elements of X (i.e. $T \subseteq \langle X \rangle$), and we will proceed by induction on n.

To this end, let $(n, g) \in T$. If $n < j_0 + m$, then clearly $(n, g) \in X$. Otherwise, assume for the inductive hypothesis that any $(p, h) \in T$ with $p < n$ is also in $\langle X \rangle$. If $n \geq j_0 + m$, then by assumption as $g \in G_n$, and furthermore as

$n - m \geq j_0$ it follows that $G_n = G_{n-m}$, so that $g \in G_{n-m}$. By the inductive hypothesis, it is assumed that $(n-m, g) \in \langle X \rangle$, and hence

$$(n-m, g)(m, 1_G) = (n, g) \in \langle X \rangle,$$

completing that proof that T is finitely generated by X. Moreover, $\mathbb{N} \times G$ has only countably many finite subsets, and hence as any subsemigroup T is finitely generated, there are at most countably many subsemigroups of $\mathbb{N} \times G$. □

This result will play an important role in determining the number of subsemigroups of $\mathbb{N} \times S$, and hence we now discuss the case where S is a union of groups. Such a semigroup is called *completely regular*, as in the following definition.

Definition 2.2.2. A semigroup S is said to be *completely regular* if there exists a family of subgroups $\{G_i : i \in \mathcal{I}\}$ of S, such that

$$S = \bigcup_{i \in \mathcal{I}} G_i.$$

That is, S is a union of groups. ■

We now give an equivalent formulation of complete regularity in terms Green's $\mathcal{H}$ relation, using Corollary 1.4.7 from Chapter 1.

Lemma 2.2.3. *A semigroup S is completely regular if and only if every $\mathcal{H}$-class of S is a group.*

Proof. ($\Leftarrow$) If every $\mathcal{H}$-class of S is a group, then noting that S is the union of its $\mathcal{H}$-classes H_s for $s \in S$ (as H_s certainly contains s), then S is a union of groups and hence is completely regular.

($\Rightarrow$) Let S be completely regular, and let $s \in S$ be arbitrary. As S is a union of groups, then s lies in some subgroup G_i of S. If e is the identity of G_i, then there exists $s' \in G_i$ such that $ss' = e = s's$.

Moreover, as $se = s = es$, then by Lemma 1.4.4, it follows that $(s, e) \in \mathcal{L}$ and

$(s, e) \in \mathcal{R}$, and hence $(s, e) \in \mathcal{H}$. Thus as $H_s = H_e$, and e is an idempotent, then by Corollary 1.4.7 it follows that H_s is a group. As $s \in S$ was chosen arbitrarily, and the set $\{H_s : s \in S\}$ covers every $\mathcal{H}$-class of S, then the result follows. □

We now give the main theorem of this section, which says the property of complete regularity of S determines the number of subsemigroups of $\mathbb{N} \times S$.

Theorem 2.2.4. *The following are equivalent for a finite semigroup S:*

(i) *$\mathbb{N} \times S$ has only countably many subsemigroups;*

(ii) *$\mathbb{N} \times S$ has only countably many pairwise non-isomorphic subsemigroups;*

(iii) *S is completely regular.*

Proof. That (i) ⇒ (ii) is immediate.

(ii) ⇒ (iii) We will prove the contrapositive. Suppose S is not completely regular. Then there exists a $\mathcal{H}$-class H of S which is not a group by Lemma 2.2.3.

Fix some $x \in H$. Then as S is finite, x has some idempotent power by Lemma 1.2.10, i.e. $x^k = x^{2k}$ for some $k \in \mathbb{N}$. It must be that $k > 1$, for otherwise x is an idempotent in H, which would then be a group by Corollary 1.4.7.

Moreover, we claim that

$$x \neq x^i \tag{2.11}$$

for any $i \geq 2$. Suppose to the contrary, that $x = x^i$ for some $i > 1$. In particular, i can be chosen so that $i > k$, as $x = x^{i^p}$ for all $p \in \mathbb{N}$. Thus

$$x = x^i = x^k x^{i-k} = x^{i-k} x^k.$$

As $k > 1$, and $x^k = x x^{k-1} = x^{k-1} x$, then by Lemma 1.4.4, it now follows that x and x^k are both $\mathcal{L}$-related and $\mathcal{R}$-related in S, and hence $\mathcal{H}$-related. But then $H = H_x = H_{x^k}$ contains the idempotent x^k, and hence H would be a group; a contradiction. Thus we have shown (2.11).

Define for $M \subseteq \mathbb{N} \setminus \{1\}$ the subsemigroup

$$T_M := \langle (1, x^k), (m, x) : m \in M \rangle \leq \mathbb{N} \times S.$$

Then the generators of T_M are all indecomposable in T_M, as 1 is indecomposable in $\mathbb{N}$ and x is indecomposable in its monogenic subsemigroup $\langle x \rangle \leq S$, for otherwise $x = x^i x^j = x^{i+j}$ for some $i, j \in \mathbb{N}$, contradicting (2.11).

We will show for $M, N \subseteq \mathbb{N} \setminus \{1\}$ with $M \neq N$ that $T_M \ncong T_N$. Suppose to the contrary, and let $\varphi : T_M \to T_N$ be an isomorphism. It can be assumed without loss of generality that $M \setminus N$ is non-empty, hence let $\hat{m} \in M \setminus N$. As x^k is idempotent, for all $m \in M$ we have

$$(1, x^k)^{mk} = (mk, (x^k)^{mk}) = (mk, x^k) = (m, x)^k,$$

and hence

$$\begin{aligned} &\varphi((1, x^k)^{mk}) = \varphi((m, x)^k) \\ \Leftrightarrow &(\varphi(1, x^k))^{mk} = (\varphi(m, x))^k. \end{aligned} \tag{2.12}$$

Applying the projection map $\pi_1 : T_N \to \mathbb{N} := (n, x^i) \mapsto n$ to (2.12) and dividing by k, for all $m \in M$ we obtain that

$$m\pi_1\varphi(1, x^k) = \pi_1\varphi(m, x), \tag{2.13}$$

and hence as $m > 1$ it follows that $\pi_1\varphi(1, x^k) < \pi_1\varphi(m, x)$. In particular, $\pi_1\varphi(m, x) > 1$ and thus $\varphi(m, x) \neq (1, x^k)$ for all $m \in M$. As $(1, x^k)$ is an indecomposable element of T_N it must be mapped onto by an indecomposable element of T_M, and hence it must be that $\varphi(1, x^k) = (1, x^k)$. But then by (2.13), it follows that $m = \pi_1\varphi(m, x)$ for all $m \in M$, and in particular that $\hat{m} = \pi_1\varphi(\hat{m}, x) \in N$ which is a contradiction.

Hence for $M, N \subseteq \mathbb{N} \setminus \{1\}$, $M \neq N$ implies that $T_M \ncong T_N$. As there are uncountably many subsets of $\mathbb{N} \setminus \{1\}$, there are uncountably many pairwise non-isomorphic subsemigroups T_M of $\mathbb{N} \times S$ as required.

(iii) $\Rightarrow$ (i) If S is completely regular, then it is a union of groups by definition, and in particular every $\mathcal{H}$-class H_s (where $s \in S$) is a group (as a subgroup G of S with identity e is contained in the $\mathcal{H}$-class H_e, which is a subgroup by Theorem 1.4.6). As

$$\mathbb{N} \times S = \mathbb{N} \times \big(\bigcup_{s \in S} H_s\big) = \bigcup_{s \in S} (\mathbb{N} \times H_s),$$

then $\mathbb{N} \times S$ is a finite union of semigroups $\mathbb{N} \times H_s$, which are pairwise disjoint. As H_s is a group for all $s \in S$, then $\mathbb{N} \times H_s$ has only countably many subsemigroups by Lemma 2.2.1. Hence it follows that $\mathbb{N} \times S$ has only countably many subsemigroups, as required. $\square$

We finish the section with some examples of finite semigroups illustrating Theorem 2.2.4

Examples 2.2.5. We will consider how many subsemigroups of $\mathbb{N} \times S$ there are up to isomorphism, for all of the two element semigroups S up to isomorphism. Namely, they are

- The two element cyclic group $\mathbb{Z}_2$ under addition modulo 2;
- The two element semilattice $\{0, 1\}$ under multiplication of real numbers;
- The two element zero semigroup $\{x, 0\}$, with multiplication defined $st = 0$ for all $s, t \in \{x, 0\}$;
- The two element left zero semigroup $\{a, b\}$, with multiplication defined $st = s$ for all $s, t \in \{a, b\}$;
- The two element right zero semigroup $\{c, d\}$, with multiplication defined $st = t$ for all $s, t \in \{c, d\}$.

For $S = \mathbb{Z}_2$, there are only countably many subsemigroups of $\mathbb{N} \times S$ by Theorem 2.2.4, as $\mathbb{Z}_2$ is a group, and hence a completely regular semigroup.

For $S = \{0, 1\}$ under multiplication, as $0^2 = 0$ and $1^2 = 1$, then both $\{0\}$ and $\{1\}$ form trivial subgroups of S, and hence as $S = \{0\} \cup \{1\}$, it is completely

regular. Hence $N \times S$ has countably many subsemigroups in this case. This is also exactly the same for $S = \{a, b\}$ (as $a^2 = a$, $b^2 = b$) and $S = \{c, d\}$.

The last case to consider is $S = \{x, 0\}$. There can be no $s \in S$ such that $sx = xs = x$, as $st = 0$ for all $s, t \in S$. Hence x cannot lie in a subgroup of S, and S is not the union of groups. Thus $\mathbb{N} \times \{x, 0\}$ has uncountably many non-isomorphic subsemigroups. By the proof of Theorem 2.2.4, the family of subsemigroups $\{T_M : M \subseteq \mathbb{N} \setminus \{1\}\}$ where T_M is given by

$$T_M = \langle (1, 0), (m, x) : m \in M \rangle = \{(n, 0), (m, x) : n \in \mathbb{N}, m \in M\}$$

is an example of uncountably many non-isomorphic subsemigroups of $\mathbb{N} \times \{x, 0\}$. △

2.3 Subdirect powers of $\mathbb{N}$

In this section, we strengthen the statement of Theorem 2.1.12 and show that of the uncountably many pairwise non-isomorphic subsemigroups of $\mathbb{N} \times \mathbb{N}$, uncountably many of them (up to isomorphism) are subdirect products. We extend this statement to the finite direct power $\mathbb{N}^k$ as in Theorem 2.1.14.

For the purposes of this section, we again adopt Notation 2.1.1. We begin by classifying when the semigroups S_M are subdirect products.

Lemma 2.3.1. *S_M is a subdirect product of $\mathbb{N} \times \mathbb{N}$ if and only if $1 \in M$.*

Proof. If S_M is a subdirect product, then there exists some $n \in \mathbb{N}$ such that $(n, 1) \in S_M$. As 1 is indecomposable in $\mathbb{N}$, then $(n, 1)$ is indecomposable in S_M, and hence belongs to X_M by Lemma 2.1.4, and so $n = 1$ and $1 \in M$.

Conversely, if $1 \in M$, then the element $(1, 1) \in S_M$. Hence for all $n \in \mathbb{N}$, the element $(1, 1)^n = (n, n) \in S_M$. Thus the projection maps from S_M onto each coordinate are surjections onto $\mathbb{N}$, and S_M is a subdirect product. □

As we saw in Theorem 2.1.12, we obtained uncountably many pairwise non-

isomorphic subsemigroups S_M of $\mathbb{N} \times \mathbb{N}$ by taking subsets M of an infinite 3-separating set. Given Lemma 2.3.1, we next establish the existence of an infinite 3-separating set which we will use to obtain uncountably many pairwise non-isomorphic subdirect products in a similar fashion.

Lemma 2.3.2. *There exists an infinite 3-separating set M_∞, with $1 \in M$.*

Proof. The proof is the same as for Corollary 2.1.11, taking M_1 to be any finite 3-separating set containing 1, such as $\{1, 2, 4\}$. □

We now give the analogy to Theorem 2.1.14, with a focus on subdirect products in the following theorem.

Theorem 2.3.3. *There are uncountably many non-isomorphic subdirect products of $\mathbb{N}^k$ for $k \geq 2$.*

Proof. Let M_∞ be an infinite 3-separating set containing 1, whose existence is established in Lemma 2.3.2. For $k = 2$, the collection

$$\mathcal{C}' = \{S_M : M \subseteq M_\infty, |M| \geq 3,\, 1 \in M\}$$

is an uncountable subset of the collection $\mathcal{C}$ established in (2.10) from Theorem 2.1.12. The proof that S_M and S_N are non-isomorphic for $M \neq N$ follows exactly as in the proof of Theorem 2.1.12, and the proof that each $S_M \in \mathcal{C}'$ is a subdirect product follows from Lemma 2.3.1.

For $k > 2$, for a subset $M \subseteq M_\infty$ with $|M| \geq 3$ and $1 \in M$, let $Y_M := \{(1, \ldots, 1, m) : m \in M\}$, and take $T_M := \langle Y_M \rangle \leq \mathbb{N}^k$. The mapping

$$\varphi : T_M \to S_M := (n, \ldots, n, p) \mapsto (n, p)$$

is verifiably an isomorphism between T_M and S_M, and that there are uncountably many such T_M up to isomorphism hence follows from the case where $k = 2$. That each T_M is a subdirect product of $\mathbb{N}^k$ follows as $(1, \ldots, 1) \in T_M$, and hence $(n, \ldots, n) \in T_M$ for all $n \in \mathbb{N}$. □

We note that in contrast to the examples stemming from the theorem of Baumslag & Roseblade given in Theorem 1.7.12, that every finitely generated commutative semigroup is finitely presented. This is a result of Rédei, which is proved in Clifford & Preston's monograph [9, Theorem 9.28]. Hence there are no examples of subsemigroups or subdirect products of $\mathbb{N}^k$ which are finitely generated, but not finitely presented. Moreover, there are no examples of finitely generated subdirect products with undecidable membership problem either. Every finitely generated subdirect product of $\mathbb{N}^k$ has decidable membership problem, as any k-tuple $(n_1, \dots n_k)$ can be decomposed into at most N generators where N is the largest of the n_i, and hence there are only finitely many possible products to check.

2.4 Subdirect products of $\mathbb{N}$ with a finite semigroup

In this section, we conclude the chapter by providing a subdirect product analogue for Theorem 2.2.4. Specifically, we will classify the finite semigroups S for which $\mathbb{N} \times S$ has only countably many subdirect products up to isomorphism. We end the section and the chapter with some examples of subdirect products of $\mathbb{N} \times S$ which are not finitely generated.

We begin by stating and proving the main result of the section.

Theorem 2.4.1. *The following are equivalent for a finite semigroup S:*

(i) *$\mathbb{N} \times S$ has only countably many subdirect products;*

(ii) *$\mathbb{N} \times S$ has only countably many pairwise non-isomorphic subdirect products;*

(iii) *For every $s \in S$, there exists some $t \in S$ such that at least one of $ts = s$ or $st = s$ holds.*

Proof. The implication (i) $\Rightarrow$ (ii) is immediate.

(ii) $\Rightarrow$ (iii) We will prove the contrapositive. Let $s \in S$ be such that

$$st \neq s \text{ and } ts \neq s \text{ for all } t \in S, \tag{2.14}$$

and fix such an s.

Suppose for a contradiction that there were some $u, t \in S$ such that $ust = s$. Then $u^n s t^n = s$ for all $n \in \mathbb{N}$. But as S is a finite semigroup, u has some idempotent power (i.e. $u^j = u^{2j}$ for some $j \in \mathbb{N}$). Hence

$$s = u^{2j} s t^{2j} = u^j s t^{2j} = s t^j$$

which contradicts (2.14). Hence we have just shown that (2.14) implies

$$ust \neq s \text{ for all } u, t \in S. \tag{2.15}$$

As S is finite, then s has some idempotent power $s^k = s^{2k}$ for some $k \in \mathbb{N}$ by Lemma 1.2.10, and in particular it must be that $k > 1$ by (2.14). Fixing such a k, we define for $M \subseteq \mathbb{N} \setminus \big(2\mathbb{N} \cup \{1\}\big)$ the semigroup

$$T_M := \big\langle (1, s^k), (2, t), (m, s) : t \in S \setminus \{s, s^k\},\ m \in M \big\rangle \leq \mathbb{N} \times S.$$

Then T_M is a subdirect product as $(1, s^k)^n = (n, s^{kn}) \in T_M$ for all $n \in \mathbb{N}$ and hence projection onto the first coordinate is surjective, and moreover every element of S appears as a second coordinate of one of the generators for T_M, so that projection onto the second coordinate is also surjective.

Next we will show that all the generators of T_M are indecomposable in T_M. Firstly, as 1 is indecomposable in $\mathbb{N}$, then $(1, s^k)$ is indecomposable in T_M. Secondly, with the given generating set, we can see that the only decomposable element in T_M of the form $(2, t)$ is $(1, s^k)^2 = (2, s^k)$, which has already been excluded from the set of generators, and hence $(2, t)$ for $t \in S \setminus \{s, s^k\}$ is indecomposable.

Finally, suppose for a contradiction that a generator of the form (m, s) were expressible as a non-trivial product of generators. Then such a product cannot include a generator of the form $(1, s^k)$ or (n, s) for $n \in M \setminus \{m\}$ by (2.14) and (2.15). But such a product can also not consist only of elements of the form $(2, t)$ because m is odd, and hence (m, s) is indecomposable for all $m \in M$. Thus the generators of T_M are indecomposable elements.

We will now show that if $M \neq N$, then $T_M \ncong T_N$. Suppose for a contradiction that $\varphi : T_M \to T_N$ were an isomorphism. Without loss of generality, we may assume $M \setminus N$ is non-empty.

For all $m \in M$, as s^k is idempotent, then

$$(1, s^k)^{mk} = (mk, (s^k)^{mk}) = (mk, s^k) = (m, s)^k,$$

and hence

$$\begin{aligned} &\varphi((1, s^k)^{mk}) = \varphi((m, s)^k) \\ \Leftrightarrow\, &(\varphi(1, s^k))^{mk} = (\varphi(m, s))^k. \end{aligned} \tag{2.16}$$

Applying the first coordinate projection map $\pi_1 : T_N \to \mathbb{N}$ to (2.16) and dividing by k gives

$$m \cdot \pi_1\varphi(1, s^k) = \pi_1\varphi(m, s) \text{ for all } m \in M, \tag{2.17}$$

and thus it follows that $\pi_1\varphi(1, s^k) < \pi_1\varphi(m, s)$ for all $m \in M$.

As $(1, s^k)$ is an indecomposable element of T_M, then the image $\varphi(1, s^k)$ must be an indecomposable element of T_N, and thus $\pi_1\varphi(1, s^k) \in \{1, 2\} \cup N$. We will in fact show that

$$\pi_1\varphi(1, s^k) = 1. \tag{2.18}$$

Having shown this, by (2.17), it follows that $m = \pi_1\varphi(m, s)$ for all $m \in M$. As $\varphi(m, s)$ is indecomposable in T_N and hence a generator, then it must be that $m \in N$ for all $m \in M$, and thus $M \subseteq N$. But this will be enough to obtain a contradiction, as $M \setminus N$ was supposed to be non-empty, thus contradicting the assumption of isomorphism.

To show (2.18), we will consider each of the cases for $\pi_1\varphi(1, s^k)$. First, suppose for a contradiction that $\pi_1\varphi(1, s^k) = 2$. Then for any $m \in M$, $\pi_1\varphi(m, s) = 2m$ by (2.17). But $\varphi(m, s)$ is an indecomposable element of T_N, and hence $\pi_1\varphi(m, s) \in \{1, 2\} \cup N$. As $N \subseteq \mathbb{N} \setminus (2\mathbb{N} \cup \{1\})$, then N consists of only odd numbers, and so it must be that $\pi_1\varphi(m, s) = 2$. This is a contradiction, as then $m = 1$ by (2.17) which implies $1 = m \in M$, but

$M \subseteq \mathbb{N} \setminus (2\mathbb{N} \cup \{1\})$.

The second case to consider for a contradiction is $\pi_1\varphi(1, s^k) \in N$. As $\pi_1\varphi(1, s^k) < \pi_1\varphi(m, s)$ for all $m \in M$ by (2.17), then $\pi_1\varphi(m, s) > 2$ for every $m \in M$. Considering the images of the generators of T_M not of the form $(2, t)$, it follows that

$$\varphi\Big(\big\{(1, s^k)\big\} \cup \big\{(m, s) \,:\, m \in M\big\}\Big) \subseteq \big\{(n, s) \,:\, n \in N\big\}. \tag{2.19}$$

As φ is an isomorphism, and the generators of both T_M and T_N are indecomposable, then the generators of the form $(2, t)$ in T_N alongside the generator $(1, s^k)$ must be mapped onto by generators of T_M. Hence by (2.19), we would have to have

$$\varphi\Big(\big\{(2, t) \,:\, t \in S \setminus \{s, s^k\}\big\}\Big) \supseteq \big\{(1, s^k)\big\} \cup \big\{(2, t) \,:\, t \in S \setminus \{s, s^k\}\big\},$$

which is impossible, as the left hand set has $|S| - 2$ elements, but the right hand set has $|S| - 1$ elements. This completes the proof of (2.18), and hence the contradiction that T_M and T_N are isomorphic for $M \neq N$.

Hence we have shown that $\big\{T_M \,:\, M \subseteq \mathbb{N} \setminus (2\mathbb{N} \cup \{1\})\big\}$ is an uncountable collection of pairwise non-isomorphic subdirect products of $\mathbb{N} \times S$.

(iii) $\Rightarrow$ (i) We will prove that every subdirect product $T \leq \mathbb{N} \times S$ is finitely generated, which will be sufficient as there are only countably many possible finite generating sets.

For every $n \in \mathbb{N}$, define the set

$$S_n := \{s \in S : (n, s) \in T\}.$$

As T is subdirect, then any $s \in S$ belongs to some S_n. Hence for every $s \in S$, choose $m_s \in \mathbb{N}$ such that $(m_s, s) \in T$, and let m be the least common multiple of all of the m_s.

We will show that

$$S_n \subseteq S_{n+m} \text{ for all } n \in \mathbb{N}. \tag{2.20}$$

Suppose $s \in S_n$, so that $(n, s) \in T$. By assumption, there exists some $t \in S$ such that either $st = s$ or $ts = s$, and in particular it follows that either $t^i s = s$ or $st^i = s$ for all $i \in \mathbb{N}$. As m is the least common multiple of all m_s, then $m = lm_t$ for some $l \in \mathbb{N}$, and we have

$$(n + m, s) = (n, s)(m_t, t)^l \in T$$

if $st = s$, and

$$(n + m, s) = (m_t, t)^l (n, s) \in T$$

if $ts = s$. In either case, $s \in S_{n+m}$ as required.

Hence every $n \in \mathbb{N}$ gives rise to an infinite ascending chain

$$S_n \subseteq S_{n+m} \subseteq S_{n+2m} \subseteq \dots$$

of subsets of S, which must eventually stabilise (meaning there exists some $i \in \mathbb{N}$ with $S_{n+im} = S_{n+jm}$ for all $j \geq i$) because S is finite. Considering the sequence $(S_i)_{i \in \mathbb{N}}$, then there must exist $j_0 \in \mathbb{N}$ such that

$$S_j = S_{j+m} \text{ for all } j \geq j_0. \tag{2.21}$$

We will show that T is generated by the finite set X, where

$$X := \bigcup_{1 \leq n < j_0 + m} \{n\} \times S_n. \tag{2.22}$$

As X is a subset of T by construction, then the subsemigroup generated by X is contained in T. Conversely, let $(n, s) \in T$. We will prove by induction on n that (n, s) can be written as a product of elements from X.

Firstly, if $n < j_0 + m$ then the element (n, s) already belongs to X, and there is nothing to show. Otherwise, suppose for the inductive hypothesis that any $(p, s') \in T$ with $p < n$ can be written as a product of elements from X. As $n \geq j_0 + m$, then $n - m \geq j_0$, and hence $S_{n-m} = S_n$ by (2.21). Hence $(n - m, s) \in T$, and thus by the inductive hypothesis $(n - m, s)$ can be written as a product of elements from X.

Recall that there exists $t \in S$ with $st = s$ or $ts = s$ (implying either $st^i = s$ or $t^i s = s$ for all $i \in \mathbb{N}$), and that we can express m as $m = lm_t$ for some $l \in \mathbb{N}$. As $m_t \leq m < j_0 + m$, then $(m_t, t) \in X$, and so we have

$$(n, s) = (n - m, s)(m_t, t)^l$$

if $st = s$, and

$$(n, s) = (m_t, t)^l (n - m, s)$$

if $ts = s$. In either case, we have expressed (n, s) as a product of elements from X. This completes the proof of finite generation of T, and hence of (iii) $\Rightarrow$ (i). $\square$

We conclude the chapter with some examples of finite semigroups illustrating Theorem 2.4.1, in comparison with Examples 2.2.5.

Examples 2.4.2. For every two element semigroup S up to isomorphism, we will consider the number of subdirect products of $\mathbb{N} \times S$ up to isomorphism.

Firstly, as we saw in Examples 2.2.5, $\mathbb{N} \times S$ has only countably many subsemigroups up to isomorphism for S being the two element cyclic group, the two element semilattice, the two element left zero semigroup and the two element right zero semigroup. Hence as subdirect products are subsemigroups, then there can also only be countably many subdirect products of $\mathbb{N} \times S$ up to isomorphism.

The only other case to consider is $S = \{x, 0\}$, the two element zero semigroup. As $st = 0$ for all $s, t \in \{x, 0\}$, however, there can be no $t \in \{x, 0\}$ such that either $xt = x$ or $tx = x$. Hence by Theorem 2.4.1, $\mathbb{N} \times S$ has uncountably many subdirect products up to isomorphism.

As in the proof of Theorem 2.4.1, for any subset $M \subseteq \mathbb{N} \setminus (2\mathbb{N} \cup \{1\})$, define

$$T_M = \langle (1, 0), (m, x) : m \in M \rangle \leq \mathbb{N} \times \mathbb{N}.$$

Then the collection $\mathcal{C} = \{T_M : M \subseteq \mathbb{N} \setminus (2\mathbb{N} \cup \{1\})\}$ is an example of an uncountable collection of pairwise non-isomorphic subdirect products of $\mathbb{N} \times \{x, 0\}$. $\triangle$

Chapter 3

Counting finitely generated subdirect products and fiber products of free semigroups

In Chapter 2, we motivated our study of subdirect products involving the free monogenic semigroup by the prior results and examples involving free groups outlined in Chapter 1 (in particular [2], [3], [13], [21]). As we saw in Chapter 2, it is perhaps surprising that the number of non-isomorphic subdirect products involving the free monogenic semigroup is uncountable. This gives some indication that even relatively basic infinite semigroups such as the free semigroups of finite rank can provide interesting and perhaps unexpected substructural behaviour.

As we noted in Chapter 1, every subdirect product of two groups arises as a fiber product of the two groups, and more generally this holds true for congruence permutable algebras due to Fleischer (Lemma 1.7.8). The varieties of semigroups and monoids are not congruence permutable however, and hence asking a question for both fiber products and subdirect products may result in inequivalent answers.

It will be an aim of Chapter 4 to discuss sufficient and necessary properties for finite generation of fiber products of free semigroups and free monoids. Hence in this chapter as a precursor, we wish to begin a combinatorial discussion

into how many such subdirect products there are, and how many of those are fiber products, similar to the aims of Chapter 2.

Hence leading on from the motivations of Chapter 2, we begin in Section 3.1 where we will define sets of "letter pairs", which can be used to naturally generate subdirect products of finitely generated semigroups. We then calculate the number of such sets of letter pairs that finitely generate a subdirect product of two free semigroups of finite rank.

In Section 3.2, we mirror the motivations and results of Section 3.1, and calculate the number of sets of letter pairs which also turn out to generate fiber products of two free semigroups of finite rank.

In Section 3.3, we finish the chapter by analytically discussing the proportion of sets of letter pairs which generate the subdirect products from Section 3.1, within the power set all possible sets of letter pairs, and comment on their abundancy. Further, we analytically discuss the proportion which also generate fiber products, within the sets of pairs that generate subdirect products, and comment on their sparsity.

We note that the contents of this chapter are largely based on the results in Section 6 of the paper [6], written by the author.

3.1 Sets of letter pairs generating subdirect products of free semigroups

In order to facilitate the investigation we outlined at the beginning of the chapter, we will adopt the following notation and definitions throughout.

Notation 3.1.1. For this chapter, A and B will be finite alphabets. A subset X of $A \times B$ will be called a *set of letter pairs*.

As sets of letter pairs are also subsets of $A^+ \times B^+$, we can consider the subsemigroups of $A^+ \times B^+$ that they generate. Hence for a given set of letter

pairs X, the maps

$$\pi_{A^+} : \langle X \rangle \to A^+ := (u, v) \mapsto u,$$
$$\pi_{B^+} : \langle X \rangle \to B^+ := (u, v) \mapsto v$$

will be projections onto the first and second coordinates respectively. The relations λ and μ will be the compositions

$$\lambda = \ker \pi_{A^+} \circ \ker \pi_{B^+},$$
$$\mu = \ker \pi_{B^+} \circ \ker \pi_{A^+},$$

recalling that the congruences $\ker \pi_{A^+}$ and $\ker \pi_{B^+}$ are defined

$$\ker \pi_{A^+} = \{((u, v), (u', v')) \in \langle X \rangle \times \langle X \rangle : u = u'\},$$
$$\ker \pi_{B^+} = \{((u, v), (u', v')) \in \langle X \rangle \times \langle X \rangle : v = v'\}.$$

Finally, for a natural number m, we will use the notation $\boldsymbol{m}$ for the set $\{1, \ldots, m\}$. ■

One intuitive way to generate a subdirect product of $A^+ \times B^+$ is by choosing a set of letter pairs X such that $\pi_A(X) = A$ and $\pi_B(X) = B$. In this way, we ensure that every $a \in A$ is paired in X with some $b \in B$ and vice versa. It follows that $\langle X \rangle$ is a subdirect product of $A^+ \times B^+$, as every $u \in A^+$ is paired in $\langle X \rangle$ with some $v \in B^+$ of the same length, and vice versa.

Conversely, if a set of letter pairs X generates a subdirect product of $A^+ \times B^+$, then for any $a \in A$, there is some $v \in B^+$ such that $(a, v) \in \langle X \rangle$. As a is an indecomposable element of A^+, then it must be that $(a, v) \in X$. Hence as a was arbitrary, then $\pi_A(X) = A$. A similar argument shows that $\pi_B(X) = B$, and hence we have proved the following.

Lemma 3.1.2. *Let X be a set of letter pairs. Then $\langle X \rangle$ is a subdirect product of $A^+ \times B^+$ if and only if $\pi_A(X) = A$ and $\pi_B(X) = B$.* □

We now illustrate this construction, with examples of sets of letter pairs that generate subdirect products of $A^+ \times B^+$.

Examples 3.1.3. (a) Let $A = B = \{a, b\}$, and $X = \{(a, a), (b, b)\}$. Then $\langle X \rangle = \{(u, u) : u \in A^+\}$ is a subdirect product of $A^+ \times B^+$, which in particular is isomorphic to A^+.

(b) Let $A = B = \{a, b\}$, and $X = \{(a, a), (b, b), (b, a)\}$. Because the only pair in X with an a as the first coordinate is (a, a), then any pair $(u, v) \in \langle X \rangle$ can have a as the i-th letter of u only if a is also the i-th letter of v. Then recalling the notation $\boldsymbol{m} = \{1, \ldots, m\}$, we can see

$$\langle X \rangle = \{(u, v) \in A^+ \times B^+ : (|u| = |v|)(\forall i \in |\boldsymbol{u}|)(|u|_i = a \Rightarrow |v|_i = a)\},$$

which is a subdirect product of $A^+ \times B^+$. △

Not all sets of letter pairs will yield a subdirect product of $A^+ \times B^+$, however, and not every subdirect product is generated by letter pairs. A natural question is to ask how many sets of letter pairs generate subdirect products of $A^+ \times B^+$. We now give the main result of this section, which gives an expression for the number of sets of letter pairs generating subdirect products of $A^+ \times B^+$.

Lemma 3.1.4. *Let A, B be finite alphabets, and further let*

$$\mathcal{S}(A \times B) := \{X \subseteq A \times B : \langle X \rangle \leq_{\text{sd}} A^+ \times B^+\}.$$

Then

$$|\mathcal{S}(A \times B)| = \sum_{k=0}^{|A|-1} (-1)^k \binom{|A|}{k} (2^{|A|-k} - 1)^{|B|}. \tag{3.1}$$

Proof. As A, B are finite, let $A = \{a_1, \ldots a_m\}$, and $B = \{b_1, \ldots b_n\}$. We will proceed by constructing a bijection between $\mathcal{S}(A \times B)$ and the set $\mathcal{M}_{m,n}(\{0, 1\})$ of $m \times n$ binary matrices with no zero rows or columns, and count the latter.

We define the mapping

$$f : \mathcal{S}(A \times B) \to \mathcal{M}_{m,n}(\{0, 1\}) := X \mapsto M_X,$$

where

$$(M_X)_{i,j} = \begin{cases} 1 & \text{if } (a_i, b_j) \in X, \\ 0 & \text{otherwise.} \end{cases}$$

We first verify that f is indeed a function. Recalling the notation $\boldsymbol{m} = \{1, \ldots, m\}$, any set $X \in \mathcal{S}(A \times B)$ has for each $i \in \boldsymbol{m}$ some $j_i \in \boldsymbol{n}$ such that $(a_i, b_{j_i}) \in X$, as $\langle X \rangle$ is subdirect. Similarly, for each $j \in \boldsymbol{n}$, there is some $i_j \in \boldsymbol{m}$ with $(a_{i_j}, b_j) \in X$. The image of any $X \in \mathcal{S}(A \times B)$ therefore has no zero rows or columns by the above argument, and is hence an element of $\mathcal{M}_{m,n}(\{0,1\})$.

Moreover, f is well defined and injective by construction. We will further show that f is surjective, and thus a bijection.

Let $M \in \mathcal{M}_{m,n}(\{0,1\})$. Consider the set

$$X = \{(a_i, b_j) \in A \times B : i \in \boldsymbol{m},\, j \in \boldsymbol{n},\, (M)_{i,j} = 1\}.$$

Then we claim $X \in \mathcal{S}(A \times B)$. As M has no zero rows, every i-th row of M has some $j_i \in \boldsymbol{n}$ such that $(M)_{i,j_i} = 1$. Hence $(a_i, b_{j_i}) \in X$ for all $i \in \boldsymbol{m}$, and thus $\pi_A(X) = A$.

Similarly, as M has no zero columns, every j-th column of M has some $i_j \in \boldsymbol{m}$ such that $(M)_{i_j,j} = 1$. Hence $(a_{i_j}, b_j) \in X$ for all $j \in \boldsymbol{n}$, and thus $\pi_B(X) = B$, proving the claim.

Moreover $f(X) = M$ by construction, and we have hence shown that f is a bijection. We now calculate $|\mathcal{M}_{m,n}(\{0,1\})|$, which will be sufficient to prove the lemma.

We proceed by an inclusion-exclusion argument. For any $i \in \boldsymbol{m}$, let A_i denote the set of binary $m \times n$ matrices with no zero columns, whose i-th row is a zero row. For a subset $I \subseteq \boldsymbol{m}$, let

$$A_I = \bigcap_{i \in I} A_i.$$

By the inclusion-exclusion principle, as $|\mathcal{M}_{m,n}(\{0,1\})|$ is precisely the num-

ber of matrices with no zero columns that do not belong to any A_i (as they can have no zero rows), then

$$|\mathcal{M}_{m,n}(\{0,1\})| = \sum_{I \subseteq \boldsymbol{m}} (-1)^{|I|} |A_I|.$$

A_I is the set of $m \times n$ binary matrices with no zero columns, whose i-th row is a zero row for each $i \in I$.

For a subset $I \subseteq \boldsymbol{m}$ and a matrix $M \in A_I$, entries in a given column of M which are not also in any of the zero rows corresponding to I can take two possible values. Excluding the singular case where all of these values are 0 (as we must exclude the zero column), then there are $2^{m-|I|} - 1$ possibilities for a fixed column of M. As there are n columns, then this gives $(2^{m-|I|} - 1)^n$ choices for M, and so $|A_I| = (2^{m-|I|} - 1)^n$. Hence

$$|\mathcal{M}_{m,n}(\{0,1\})| = \sum_{I \subseteq \boldsymbol{m}} (-1)^{|I|} (2^{m-|I|} - 1)^n.$$

Noting that as the summands depend only on $|I|$ which varies from 0 to m, and there are $\binom{m}{|I|}$ possible subsets of size $|I|$, then we have

$$|\mathcal{M}_{m,n}(\{0,1\})| = \sum_{|I|=0}^{m} (-1)^{|I|} \binom{m}{|I|} (2^{m-|I|} - 1)^n.$$

Changing summation index to k and noting that the $k = m$ summand is equal to zero, then we have shown that

$$|\mathcal{M}_{m,n}(\{0,1\})| = \sum_{k=0}^{m-1} (-1)^k \binom{m}{k} (2^{m-k} - 1)^n.$$

We have now shown (3.1), as $m = |A|$, $n = |B|$, and $|\mathcal{M}_{m,n}(\{0,1\})| = |\mathcal{S}(A \times B)|$, thus completing the proof of the lemma. $\square$

3.2 Sets of letter pairs generating fiber products of free semigroups

In this section, we now ask how many sets of letter pairs generate fiber products of $A^+ \times B^+$. We briefly recount Fleischer's lemma for the reader, which characterises when a subdirect product is a fiber product.

Lemma 1.7.8 (Fleischer's lemma, [4, Lemma 10.1])**.** *Let S, T, U be semigroups, and let $U \leq_{\mathrm{sd}} S \times T$. For the projection maps*

$$\pi_S : U \to S := (s,t) \mapsto s,$$
$$\pi_T : U \to T := (s,t) \mapsto t,$$

denote by σ the congruence $\ker \pi_S$ on U, and denote by ρ the congruence $\ker \pi_T$ on U. Then U is a fiber product of S with T if and only if

$$\sigma \circ \rho = \rho \circ \sigma.$$

Recalling Notation 3.1.1, this is equivalent to $\lambda = \mu$. We now use this lemma in the following examples.

Examples 3.2.1. (a) We saw in Examples 3.1.3 (a) that for $A = B = \{a, b\}$, the subsemigroup

$$\langle X \rangle = \{(u,u) : u \in A^+\}$$

of $A^+ \times B^+$ is a subdirect product generated by the set of letter pairs $X = \{(a,a),(b,b)\}$. Recalling Notation 3.1.1, for two elements $(u,u),(v,v) \in \langle X \rangle$ we have

$$\begin{aligned}((u,u),(v,v)) \in \lambda &\Leftrightarrow (u,v) \in \langle X \rangle \\ &\Leftrightarrow u = v \\ &\Leftrightarrow (v,u) \in \langle X \rangle \\ &\Leftrightarrow ((u,u),(v,v)) \in \mu.\end{aligned}$$

Hence $\lambda = \mu$, and $\langle X \rangle$ is a fiber product of $A^+ \times B^+$ by Fleischer's lemma.

(b) Recalling Examples 3.1.3 (b), we saw that for $A = B = \{a, b\}$, the subsemigroup

$$\langle X \rangle = \{(u, v) \in A^+ \times B^+ : (|u| = |v|)(\forall i \in |\boldsymbol{u}|)(|u|_i = a \Rightarrow |v|_i = a)\},$$

is a subdirect product of $A^+ \times B^+$ generated by the set of letter pairs $X = \{(a, a), (b, b), (b, a)\}$. Recalling Notation 3.1.1, we have $((a, a), (b, b)) \in \mu$, as

$$((a, a), (b, b)) \in \mu \Leftrightarrow (b, a) \in \langle X \rangle,$$

but $((a, a), (b, b)) \notin \lambda$, as

$$((a, a), (b, b)) \in \lambda \Leftrightarrow (a, b) \in \langle X \rangle.$$

Hence as $\lambda \neq \mu$, then $\langle X \rangle$ is not a fibered product of $A^+ \times B^+$ by Fleischer's lemma. △

We will shortly establish a bijection between the sets of letter pairs generating fiber products of $A^+ \times B^+$, and a certain set of binary $|A| \times |B|$ matrices. In order to do this, we first remind the reader of the definition of a *submatrix*.

Definition 3.2.2. Let $m \times n$ matrix M with entries a_{ij} for $i \in \boldsymbol{m}$, $j \in \boldsymbol{n}$. A *submatrix* M' of M is an $|I| \times |J|$ matrix with entries a_{ij} for $i \in I$, $j \in J$, where I is some non-empty subset of $\boldsymbol{m}$, and J is some non-empty subset of $\boldsymbol{n}$. ■

We now establish the aformentioned bijection in the following result.

Lemma 3.2.3. *Let A, B be finite alphabets, and further let*

$$\mathcal{F}(A \times B) = \{X \subseteq A \times B : \langle X \rangle \leq_{\text{fp}} A^+ \times B^+\}.$$

Then $\mathcal{F}(A \times B)$ is in bijection with the set $\mathcal{M}^*_{|A|,|B|}(\{0, 1\})$ of $|A| \times |B|$ binary

matrices with no zero rows and columns, that do not contain any of

$$\begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \tag{3.2}$$

as submatrices.

Proof. Let $A = \{a_1, \ldots, a_m\}$, $B = \{b_1, \ldots, b_m\}$.

Define the mapping

$$f : \mathcal{F}(A \times B) \to \mathcal{M}^*_{m,n}(\{0,1\}) := X \mapsto M_X,$$

where

$$(M_X)_{i,j} = \begin{cases} 1 & \text{if } (a_i, b_j) \in X, \\ 0 & \text{otherwise.} \end{cases}$$

We will first verify that f is a function with codomain $\mathcal{M}^*_{m,n}(\{0,1\})$. As in the proof of Lemma 3.1.4, the image of every set $X \in \mathcal{F}(A \times B)$ certainly has no zero rows or columns, as $\langle X \rangle$ is subdirect. Moreover, suppose for a contradiction that the image of some $X \in \mathcal{F}(A \times B)$ contains the submatrix

$$\begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}.$$

Then by the definition of f, there exist indices $i_1, i_2 \in \boldsymbol{m}$ and $j_1, j_2 \in \boldsymbol{n}$ where $(a_{i_1}, b_{j_1}) \notin X$, but $(a_{i_1}, b_{j_2}), (a_{i_2}, b_{j_1}), (a_{i_2}, b_{j_2}) \in X$. Recalling Notation 3.1.1, it follows that $((a_{i_1}, b_{j_2}), (a_{i_2}, b_{j_1})) \in \mu$ as

$$((a_{i_1}, b_{j_2}), (a_{i_2}, b_{j_1})) \in \mu \Leftrightarrow (a_{i_2}, b_{j_2}) \in X,$$

but $((a_{i_1}, b_{j_2}), (a_{i_2}, b_{j_1})) \notin \lambda$, as

$$((a_{i_1}, b_{j_2}), (a_{i_2}, b_{j_1})) \in \lambda \Leftrightarrow (a_{i_1}, b_{j_1}) \in X.$$

This is a contradiction of Fleischer's lemma, as $\langle X \rangle \leq_{\text{fp}} A^+ \times B^+$, but $\lambda \neq \mu$. Similar contradictions will be obtained supposing the image of some $X \in \mathcal{F}(A \times B)$ contains any other submatrix from (3.2), and hence f is a

function with codomain $\mathcal{M}^*_{m,n}(\{0,1\})$.

Again f is well-defined and injective by construction, hence it remains to show that f is surjective. For a matrix $M \in \mathcal{M}^*_{i,j}(\{0,1\})$, let

$$X = \{(a_i, b_j) \in A \times B : i \in \boldsymbol{m},\, j \in \boldsymbol{n},\, (M)_{i,j} = 1\}.$$

Then we claim $\langle X \rangle$ is a fiber product of $A^+ \times B^+$, and hence that $X \in \mathcal{F}(A \times B)$. The argument that X is a subdirect product follows exactly as in the proof of Lemma 3.1.4. Assume for a contradiction that $\langle X \rangle$ is not a fiber product. By Fleischer's lemma, there exist two pairs $(u, v), (u', v') \in \langle X \rangle$ such that either

$$((u, v), (u', v')) \in \lambda \text{ and } ((u, v), (u', v')) \notin \mu,$$

or vice versa. Note for two pairs $(u, v), (u', v') \in \langle X \rangle$ that

$$\begin{aligned} ((u, v), (u', v')) \in \lambda &\Leftrightarrow (u, v') \in \langle X \rangle \\ &\Leftrightarrow \forall i \in |\boldsymbol{u}|,\, (|u|_i, |v'|_i) \in X \\ &\Leftrightarrow \forall i \in |\boldsymbol{u}|,\, ((|u|_i, |v|_i), (|u'|_i, |v'|_i)) \in \lambda, \end{aligned}$$

and a similar proof also gives that

$$((u, v), (u', v')) \in \mu \Leftrightarrow \forall i \in |\boldsymbol{u}|,\, ((|u|_i, |v|_i), (|u'|_i, |v'|_i)) \in \mu.$$

Hence there are two pairs $(a_{i_1}, b_{j_1}), (a_{i_2}, b_{j_2}) \in X$ with either

$$((a_{i_1}, b_{j_1}), (a_{i_2}, b_{j_2})) \in \lambda \text{ and } ((a_{i_1}, b_{j_1}), (a_{i_2}, b_{j_2})) \notin \mu,$$

or vice versa. Thus either

$$(a_{i_1}, b_{j_2}) \in X \text{ and } (a_{i_2}, b_{j_1}) \notin X,$$

or vice versa. In either case, it must be that $i_1 \neq i_2$ and $j_1 \neq j_2$ to avoid contradiction, and hence the matrix entries $(M)_{i_1,j_1}, (M)_{i_1,j_2}, (M)_{i_2,j_1}$ and $(M)_{i_2,j_2}$ form the corners of a 2×2 submatrix M' of M. But now as ex-

actly three of $(a_{i_1}, b_{j_1}), (a_{i_2}, b_{j_2}), (a_{i_1}, b_{j_2}), (a_{i_2}, b_{j_1})$ are in X, M' must be a submatrix from (3.2), which is a contradiction of choice of M.

Hence $\langle X \rangle$ is a fiber product, and thus $X \in \mathcal{F}(A \times B)$. By construction, $f(X) = M$, and thus we conclude the proof that f is surjective, and hence a bijection. □

As a corollary of this result, we will count the number of sets of letter pairs that generate fiber products of $A^+ \times B^+$. We first give a necessary definition used in the proof of this result, which utilises the proof of [17, Theorem 3.1]. As such, note that we now drop Notation 3.1.1, so that our proof coincides with the notation of [17, Theorem 3.1].

Definition 3.2.4. For $k \in \mathbb{N}$, a *permutation σ of length k* is a length k string over the alphabet $\boldsymbol{k}$, containing each element of $\boldsymbol{k}$ exactly once. The set of all permutations of length k is denoted S_k. For example, 3214 is a permutation in S_4.

Given a pemutation σ of length k, the i-th *letter* of σ is the i-th element of the string. This is denoted $|\sigma|_i$. For example, $|3214|_3 = 1$. ■

We now give the main result of the section.

Corollary 3.2.5. *Let A, B be finite alphabets, and further let*

$$\mathcal{F}(A \times B) = \{X \subseteq A \times B : \langle X \rangle \leq_{\text{fp}} A^+ \times B^+\}.$$

Then

$$|\mathcal{F}(A \times B)| = \sum_{k=1}^{\min\{|A|,|B|\}} k!\, S_2(|A|, k)\, S_2(|B|, k), \tag{3.3}$$

where $S_2(n, k)$ is the Stirling number of the second kind.

Proof. As Lemma 3.2.3 establishes a bijection between $\mathcal{F}(A \times B)$ and the set $\mathcal{M}^*_{|A|,|B|}(\{0, 1\})$ of $|A| \times |B|$ binary matrices with no zero rows and columns

that do not contain any of

$$\begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$$

as submatrices, then it suffices to calculate $|\mathcal{M}^*_{|A|,|B|}(\{0,1\})|$. We will utilise and follow the proof of [17, Theorem 3.1], which calculates the number of $|A| \times |B|$ binary matrices with no submatrices from (3.2) to be

$$\sum_{k=0}^{\min\{|A|,|B|\}} k!\, S_2(|A|+1, k+1)\, S_2(|B|+1, k+1).$$

We will establish a bijection between $\mathcal{M}^*_{|A|\times|B|}(\{0,1\})$ and the set of all triples (μ, ν, σ) where μ, ν and σ are as follows for some $1 \leq k \leq \min\{|A|, |B|\}$:

(i) μ, ν are set partitions of $\boldsymbol{|A|}$ and $\boldsymbol{|B|}$ respectively into k subsets;

(ii) σ is a permutation of length k (recalling Definition 3.2.4).

As the authors of [17, Theorem 3.1] comment, any matrix $M \in \mathcal{M}^*_{|A|\times|B|}(\{0,1\})$ can be transformed via row and column exchanges into a block diagonal, $|A| \times |B|$ matrix of the form

$$\begin{pmatrix} [1] & [0] & \cdots & [0] & [0'] \\ [0] & [1] & \cdots & [0] & [0'] \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ [0] & [0] & \cdots & [1] & [0'] \\ [0'] & [0'] & \cdots & [0'] & [0'] \end{pmatrix}$$

where [1] is a block of ones, [0] is a block of zeros, and [0′] is either a block of zeros or empty. For any $|A| \times |B|$ binary matrix with no submatrices from (3.2), after this transformation, the diagonal consists of blocks of ones, except possibly the last block which could be a block of zeros. As M has no zero rows and columns however, this is not a possibility. Hence any matrix $M \in \mathcal{M}^*_{|A|\times|B|}(\{0,1\})$ can be transformed via row and column exchanges into

a $|A| \times |B|$ block diagonal matrix of the form

$$\begin{pmatrix} [1] & [0] & \cdots & [0] \\ [0] & [1] & \cdots & [0] \\ \vdots & \vdots & \ddots & \vdots \\ [0] & [0] & \cdots & [1] \end{pmatrix}.$$

Having originally labelled the rows and columns of M by 1 to $|A|$ and 1 to $|B|$ respectively in the natural way, we can track the resulting effects of the row and column operations, and then associate the [1] blocks with the permuted row and column labels as in [17, Theorem 3.1]. We then obtain two set partitions $\mu = \{C_1, \ldots, C_k\}$ and $\nu = \{D_1, \ldots, D_k\}$, where the subsets of μ are ordered by the largest element, and likewise for ν.

Moreover, in exactly the same way as [17, Theorem 3.1], we also obtain a length k permutation σ from this process, where σ is defined by

$$|\sigma|_i = j \Leftrightarrow C_i \text{ and } D_j \text{ form a [1]-block.}$$

This hence associates a unique triple (μ, ν, σ) to any given $M \in \mathcal{M}^*_{|A|,|B|}(\{0,1\})$. Conversely, given a triple (μ, ν, σ) where $\mu = \{C_1, \ldots, C_k\}$ is ordered by the largest element of each subset, $\nu = \{D_1, \ldots, D_k\}$ likewise, and σ is a length k permutation, then we can define the matrix M where

$$(M)_{i,j} = \begin{cases} 1 & \text{if } (i,j) \in C_l \times D_{|\sigma|_l} \text{ for some } l \in \boldsymbol{k}, \\ 0 & \text{otherwise.} \end{cases}$$

Suppose for a contradiction that $(M)_{i,j}$ contains the submatrix

$$\begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}.$$

Then there are indices $i_1, i_2 \in \boldsymbol{|A|}$, $j_1, j_2 \in \boldsymbol{|B|}$ such that

$$(i_2, j_1) \in C_p \times D_{|\sigma|_p} \text{ for some } p \in \boldsymbol{k}, \tag{3.4}$$

$$(i_1, j_2) \in C_q \times D_{|\sigma|_q} \text{ for some } q \in \boldsymbol{k}, \tag{3.5}$$

$$(i_2, j_2) \in C_r \times D_{|\sigma|_r} \text{ for some } r \in \boldsymbol{k}, \tag{3.6}$$

but

$$(i_1, j_1) \notin C_l \times D_{|\sigma|_l} \text{ for any } l \in \boldsymbol{k}. \tag{3.7}$$

As μ is a partition, then it must follow that $p = r$ from (3.4) and (3.6). Similarly as ν is a partition and σ is a permutation, then it must also follow that $q = r$ from (3.5) and (3.6), and so $p = q = r$. But then as $i_1 \in C_p$ and $j_1 \in D_{|\sigma|_p}$ by (3.5) and (3.4), then $(i_1, j_1) \in C_p \times D_{|\sigma|_p}$, which contradicts (3.7).

Similar contradictions will be obtained assuming that $(M)_{i,j}$ contains any of the other submatrices from (3.2). Moreover, as any $i \in |\boldsymbol{A}|$ belongs to some C_l as μ is a partition, then there is an index $j \in |\boldsymbol{B}|$ such that $(i, j) \in C_l \times D_{|\sigma|_l}$. Hence M has no zero rows by construction. Similarly, M has no zero columns, and hence $M \in \mathcal{M}^*_{|A|,|B|}(\{0, 1\})$ as required.

This completes the proof that $\mathcal{M}^*_{|A|,|B|}(\{0, 1\})$ is in bijection with the set of all described triples (μ, ν, σ). There are $S_2(|A|, k)$ partitions of $|\boldsymbol{A}|$ into k subsets, $S_2(|B|, k)$ partitions of $|\boldsymbol{B}|$ into k subsets, and $k!$ permutations of length k. As k can be at most $\min\{|A|, |B|\}$, then we have proved (3.3). □

3.3 Proportion of sets of letter pairs generating subdirect products and fiber products

In this section, we want to use the results obtained in Lemma 3.1.4 to consider the proportion of sets of letter pairs that generate subdirect products of $A^+ \times B^+$, within the power set of sets of all letter pairs, when allowing $|A|$ and $|B|$ to grow arbitrarily large, but be equal to each other. That is, determining

$$\lim_{|A|,|B| \to \infty} \frac{|\mathcal{S}(A \times B)|}{|\mathcal{P}(A \times B)|}$$

with $|A| = |B|$. We hence consider the proportion of all generating sets of letter pairs which give a subdirect product of $A^+ \times A^+$, letting $|A|$ grow in a limit. We now introduce the next result to comment that there are

an abundance of sets of letter pairs $X \in \mathcal{P}(A \times A)$ which give subdirect products.

Proposition 3.3.1. *For a given finite alphabet $A = \{a_1, \ldots, a_m\}$, let*

$$\mathcal{S}(A \times A) := \{X \subseteq A \times A : \langle X \rangle \leq_{\text{sd}} A^+ \times A^+\}.$$

Then

$$\lim_{|A| \to \infty} \frac{|\mathcal{S}(A \times A)|}{|\mathcal{P}(A \times A)|} = 1. \tag{3.8}$$

Proof. Clearly as $\mathcal{S}(A \times A) \subseteq \mathcal{P}(A \times A)$, then

$$\frac{|\mathcal{S}(A \times A)|}{|\mathcal{P}(A \times A)|} \leq 1.$$

We will now utilise Lemma 3.1.4 to obtain a lower bound for the above ratio, and show that the limit of this lower bound is 1, proving (3.8) by the sandwich theorem. Recall that in Lemma 3.1.4, we found that

$$|\mathcal{S}(A \times B)| = \sum_{k=0}^{|A|-1} (-1)^k \binom{|A|}{k} (2^{|A|-k} - 1)^{|B|}. \tag{3.9}$$

As we will let $|A|$ grow arbitrarily large in the limit, we can assume $|A| > 2$ without loss of generality. Hence considering $|\mathcal{S}(A \times A)|$, the sum of the first two terms ($k = 0$ and $k = 1$) from (3.9) is precisely

$$(2^{|A|} - 1)^{|A|} - |A|(2^{|A|-1} - 1)^{|A|}.$$

For the remaining terms ($k \geq 2$) in (3.1), we note that the sequence

$$(x_k)_{k=2}^{|A|-1} \text{ where } x_k := \binom{|A|}{k} (2^{|A|-k} - 1)^{|A|}$$

is strictly decreasing, as

$$\frac{x_i}{x_{i+1}} = \frac{i+1}{|A|-i} \left(2 + \frac{1}{2^{|A|-(i+1)} - 1}\right)^{|A|} > \frac{2^{|A|}}{|A|} > 1$$

for $i \in \{2, \ldots, |A| - 2\}$. Hence

$$\sum_{k=2}^{|A|-1} (-1)^k \binom{|A|}{k} (2^{|A|-k} - 1)^{|A|} \geq 0,$$

and thus

$$(2^{|A|} - 1)^{|A|} - |A|(2^{|A|-1} - 1)^{|A|} < |\mathcal{S}(A \times A)|. \tag{3.10}$$

Hence as

$$\begin{aligned}\left(1 - \frac{1}{2^{|A|}}\right)^{|A|} - \frac{|A|}{2^{|A|}} &= \frac{(2^{|A|} - 1)^{|A|} - |A|(2^{|A|-1})^{|A|}}{2^{|A|^2}} \\ &\leq \frac{(2^{|A|} - 1)^{|A|} - |A|(2^{|A|-1} - 1)^{|A|}}{2^{|A|^2}},\end{aligned}$$

then combining the above and (3.10) gives

$$\left(1 - \frac{1}{2^{|A|}}\right)^{|A|} - \frac{|A|}{2^{|A|}} \leq \frac{|\mathcal{S}(A \times A)|}{|\mathcal{P}(A \times A)|} \leq 1.$$

We will have completed the proof of the proposition having shown

$$\lim_{|A|\to\infty} \left(\left(1 - \frac{1}{2^{|A|}}\right)^{|A|} - \frac{|A|}{2^{|A|}} \right) = \lim_{|A|\to\infty} \left(1 - \frac{1}{2^{|A|}}\right)^{|A|} = 1,$$

which is equivalent to showing

$$\lim_{|A|\to\infty} \log \left(1 - \frac{1}{2^{|A|}}\right)^{|A|} = 0. \tag{3.11}$$

As the limit in (3.11) is equivalent to

$$\lim_{|A|\to\infty} \frac{\log\left(1 - 2^{-|A|}\right)}{|A|^{-1}},$$

then using L'Hôpital's rule gives

$$\begin{aligned}\lim_{|A|\to\infty} \frac{\log\left(1 - 2^{-|A|}\right)}{|A|^{-1}} &= \lim_{|A|\to\infty} \frac{2^{-|A|}\log(2)}{-|A|^{-2}(1 - 2^{-|A|})} \\ &= \lim_{|A|\to\infty} \frac{-|A|^2}{2^{|A|}} \cdot \frac{\log(2)}{(1 - 2^{-|A|})}\end{aligned}$$

which is zero, as $|A|^2$ grows asymptotically slower than $2^{|A|}$, hence proving (3.11) and thus the proposition. □

Similarly to Proposition 3.3.1, we now consider the proportion of sets of letter pairs $X \in \mathcal{S}(A \times B)$ that also generate fiber products of $A^+ \times B^+$, allowing $|A|$ and $|B|$ to grow arbitrarily large, but be equal to each other. In the following result, even though generating sets for subdirect products are abundant, we will see that the same is not true for fiber products.

In particular, we now show in the final result of the chapter that almost no sets of letter pairs generate fiber products of $A^+ \times A^+$, letting $|A|$ and grow in a limit.

Proposition 3.3.2. *For a given finite alphabet $A = \{a_1, \ldots, a_m\}$, let*

$$\mathcal{S}(A \times A) := \{X \subseteq A \times A : \langle X \rangle \leq_{\text{sd}} A^+ \times A^+\},$$

and let

$$\mathcal{F}(A \times A) = \{X \subseteq A \times A : \langle X \rangle \leq_{\text{fp}} A^+ \times A^+\}.$$

Then

$$\lim_{|A| \to \infty} \frac{|\mathcal{F}(A \times A)|}{|\mathcal{S}(A \times A)|} = 0. \tag{3.12}$$

Proof. By Corollary 3.2.5,

$$|\mathcal{F}(A \times A)| = \sum_{k=1}^{|A|} k!\, S_2(|A|, k)^2 \tag{3.13}$$

where $S_2(|A|, k)$ is the Stirling number of the second kind. The number of ways to partition the set $\boldsymbol{|A|}$ into k non-empty unlabelled blocks is less than the number of ways to assign the elements of $\boldsymbol{|A|}$ into k possibly empty unlabelled blocks. There are $S_2(|A|, k)$ ways to achieve the former, and $\frac{k^{|A|}}{k!}$ ways to achieve the latter. Hence

$$|\mathcal{F}(A \times A)| \leq \sum_{k=1}^{|A|} k! \left(\frac{k^{|A|}}{k!}\right)^2 \leq \sum_{k=1}^{|A|} k^{2|A|} \leq |A| \cdot |A|^{2|A|} \leq |A|^{3|A|}. \tag{3.14}$$

Now utilising Proposition 3.3.1 and (3.14), we have

$$0 \leq \lim_{|A|\to\infty} \frac{|\mathcal{F}(A\times A)|}{|\mathcal{S}(A\times A)|} = \lim_{|A|\to\infty} \frac{|\mathcal{F}(A\times A)|}{|\mathcal{P}(A\times A)|} \leq \lim_{|A|\to\infty} \frac{|A|^{3|A|}}{2^{|A|^2}}.$$

As

$$\lim_{|A|\to\infty} \frac{|A|^{3|A|}}{2^{|A|^2}} = \lim_{|A|\to\infty} (2^{-|A|})^{(|A|-3\log_2(|A|))} \tag{3.15}$$

and $\lim_{|A|\to\infty}(|A| - 3\log_2(|A|)) = \infty$ by standard analytic arguments, then (3.15) is equal to zero, and (3.12) follows. □

Chapter 4

Finitary properties for fiber products of free semigroups and monoids

As discussed in Chapter 2 and Chapter 3, our consideration for subdirect products of free semigroups in particular has been motivated by prior results and examples involving free groups outlined in Chapter 1. We also noted that subdirect products and fiber products only coincide in congruence permutable varieties by Fleischer (Lemma 1.7.8), and hence not in the varieties of monoids and semigroups. Hence the classical questions relating to finitary properties that have been asked for subdirect products of groups can be asked for fiber products of monoids and semigroups, with possibly inequivalent answers.

In the case of subdirect products of free groups, we outlined some of these finitary properties such as finite generation, finite presentability, decidability of the membership problem, word problem, and so on that have been well studied (see Section 1.7). The literature for semigroups and monoids related to finitary properties is more recent however, as the area is somewhat less explored.

As subdirect products have a general setting in universal algebra, this allows for the natural finitary property questions of decidability, generation and presentation to be asked that were answered for free groups. Mayr &

Ruškuc [19] have investigated such questions of generation and presentation for subdirect products of general algebras. Specifically relevent to this work, via congruences on a free monoid, they gave an example of a fiber product over a finitely presented quotient which is itself finitely presented, but whose factors are not finitely presented [19, Example 7.3].

They also gave an example of a fiber product of finitely generated free monoids over a finite fiber which is not finitely generated [19, Example 7.1], which motivated them to ask the following question:

Question 4.0.1 ([19, Problem 7.2])**.** *Find necessary and sufficient conditions for a fiber product of finitely generated monoids over a finite monoid to be finitely generated. More specifically, is it decidable whether a fiber product of two finitely generated free monoids over a finite quotient is finitely generated?*

In this chapter, we take the points outlined above (and our findings from Chapter 2 and Chapter 3) as our motivation to begin a study of finitary properties for fiber products of free semigroups and monoids.

We begin in Section 4.1, where we will concentrate on finite generation for fiber products of free monoids. Namely, we aim to show that finite generation of a fiber product of two (finitely generated) free monoids over a finite fiber is decidable. We will also determine all such finite fibers ensuring finite generation, directly answering Question 4.0.1. Moreover in Section 4.2, we will show that finite generation implies finite presentation for these fiber products, as well as giving an appropriate presentation.

In Section 4.3 we consider the same question, this time for free semigroups, and, perhaps somewhat surprisingly, show that the answers in two cases differ substantially. In particular for finite fibers, we will show that there are no finitely generated fiber products of two free semigroups. In the more general case, we will show that finitely generated fiber products of free semigroups must have finitely generated, $\mathcal{J}$-trivial, idempotent-free fiber quotients. Given this, we will show that these conditions are not sufficient for finite generation by considering the case where the fiber quotient is a free

commutative semigroup.

We note that the results of this section are based on Sections 3 and 4 of the paper [6], written by the author.

4.1 Finite generation for fiber products of free monoids over finite fibers

In Chapter 3, we saw for two finite alphabets A and B that a set of letter pairs is likely to finitely generate a subdirect product of $A^+ \times B^+$ (Proposition 3.3.1). We also saw that the same cannot be said for generating a fiber product of $A^+ \times B^+$ (Proposition 3.3.2). Alongside these results, it is with the motivation of Question 4.0.1 that in this section, we will investigate finite generation for fiber products, particularly for free monoids.

It will be our main aim to classify for which finite monoid fibers M and which epimorphisms $\varphi : A^* \to M$, $\psi : B^* \to M$ is the fiber product of A^* with B^* with respect to φ, ψ finitely generated. We hence begin the section by briefly recalling Definition 1.7.5 for the reader.

Definition 1.7.5. Given semigroups S, T, U and epimorphisms $\varphi : S \to U$, $\psi : T \to U$, the *fiber product of S and T with respect to φ, ψ* is the set

$$\Pi(\varphi, \psi) := \{(s, t) \in S \times T : \varphi(s) = \psi(t)\}$$

with multiplication inherited from $S \times T$. U is called the *fiber*, or *fiber quotient* of $\Pi(\varphi, \psi)$. If V is a subdirect product of $S \times T$ which is also a fiber product, we will write $V \leq_{\text{fp}} S \times T$. ■

The notation from Definition 1.7.5 will be adopted throughout this section. We will also focus on when the alphabets A and B are finite, as a consequence of the following result.

Proposition 4.1.1. *Let S, T and U be semigroups, and let $\varphi : S \to U$, $\psi : T \to U$ be epimorphisms. If $\Pi(\varphi, \psi)$ is finitely generated as a semigroup, then so are S, T and U.*

Proof. The projection maps from any generating set X of $\Pi(\varphi, \psi)$ are semigroup epimorphisms onto S and T, and hence as X is a finite generating set for $\Pi(\varphi, \psi)$, then the projection map images are finite generating sets for S and T by Lemma 1.6.10. As φ is a semigroup epimorphism from S to U, and S is finitely generated, then the image of any finite generating set for S is a finite generating set for U, again by Lemma 1.6.10. □

We note that a free monoid A^* over an alphabet A is finitely generated as a monoid if and only if the free semigroup A^+ over A is finitely generated as a semigroup, which is precisely when A is finite. Thus throughout the section, A and B will be finite alphabets.

Our next result shows that for a finitely generated fiber product of two free monoids over a finite fiber, the associated fiber need be a group.

Lemma 4.1.2. *For two finite alphabets A and B, let $\varphi : A^* \to M$ and $\psi : B^* \to M$ be epimorphisms onto a finite monoid M. If $\Pi(\varphi, \psi)$ is finitely generated, then M is a group.*

Proof. We will prove the contrapositive: if M is not a group, then $\Pi(\varphi, \psi)$ is not finitely generated.

Recalling Lemma 1.2.10 and Lemma 1.2.11, as M is assumed to be a finite non-group, then there exists some $m \in M$ and $k \in \mathbb{N}$ such that m^k is idempotent but m^k is not the identity of the monoid. Consequently, m has no inverse in M. In particular, m can be chosen from $\varphi(A)$ which is a generating set for M by Lemma 1.6.9, for otherwise every generator from $\varphi(A)$ would have an inverse and M would be a group.

Hence as ψ is also a surjection, there exists some $a \in A$, $v \in B^*$ such that $\varphi(a) = m = \psi(v)$. It now follows that the set

$$\{(a^{ik}, v^k) : i \in \mathbb{N}\} \tag{4.1}$$

is contained in $\Pi(\varphi, \psi)$, as

$$\varphi(a^{ik}) = \varphi(a)^{ik} = m^{ik} = (m^k)^i = m^k = (\psi(v))^k = \psi(v^k)$$

for all $i \in \mathbb{N}$. Suppose for a contradiction that $\Pi(\varphi, \psi)$ were finitely generated, and let

$$X = \{(u_i, v_i) : 1 \leq i \leq p\} \subseteq A^* \times B^* \tag{4.2}$$

be a finite generating set. We will attempt to write any element (a^{ik}, v^k) from (4.1) as a finite product of elements of X.

As $\varphi(a^j) = \varphi(a)^j = m^j$ for all $j \in \mathbb{N}$, it follows that $\varphi(a^j) \neq 1_M$ for all $j \in \mathbb{N}$, as m was assumed to have no inverse. Hence denoting the empty word of B^* by ε_B (which maps to 1_M under ψ necessarily), it follows that $(a^j, \varepsilon_B) \notin \Pi(\varphi, \psi)$ for all $j \in \mathbb{N}$. Thus any element (a^{ik}, v^k) from (4.1) can be written as a finite product of elements of X', where

$$X' = \{(u_i, v_i) \in X : v_i \neq \varepsilon_B\}.$$

The element (a^{ik}, v^k) can be decomposed into at most $k|v|$ elements of X', as v^k can be decomposed into at most $k|v|$ elements of B^+. As

$$(a^{ik}, v^k) = (u_{i_1}, v_{i_1})(u_{i_2}, v_{i_2}) \dots (u_{i_q}, v_{i_q})$$

for some elements $(u_{i_j}, v_{i_j}) \in X'$ with $q \leq k|v|$, then it follows that

$$|a^{ik}| \leq lk|v|,$$

where $l = \max_{1 \leq i \leq p} |u_i|$. This is a contradiction, as the above inequality must hold for all $i \in \mathbb{N}$. Hence $\Pi(\varphi, \psi)$ cannot be finitely generated. □

In the next result, we refine the condition from Lemma 4.1.2, and show that for a finitely generated fiber product of two free monoids with a finite fiber, the associated fiber need be a cyclic group.

Lemma 4.1.3. *For two finite alphabets A and B, let $\varphi : A^* \to G$ and $\psi : B^* \to G$ be epimorphisms onto a finite group G. If $\Pi(\varphi, \psi)$ is finitely*

generated, then G is cyclic.

Proof. We will prove the contrapositive: if G is not cyclic, then $\Pi(\varphi, \psi)$ is not finitely generated. We will proceed by creating an infinite set of indecomposable elements for $\Pi(\varphi, \psi)$. The result will then follow, as any decomposable element of $\Pi(\varphi, \psi)$ must be included in a generating set.

By Lemma 1.6.9, every element of $G \setminus \{1_G\}$ is expressible as a finite product of elements from $\varphi(A)$. Hence $\varphi(A)$ is a monoid generating set for G, and thus in particular also a group generating set.

As G is non-cyclic, we claim there exist two distinct elements $g, h \in \varphi(A)$ such that

$$gh^i \neq 1_G \text{ for all } i \in \mathbb{N}^0. \tag{4.3}$$

For otherwise, we can choose $x \in \varphi(A)$ such that for all $g \in \varphi(A)$ there is some $i \in \mathbb{N}^0$ with $gx^i = 1_G$. This implies $g = x^{-i}$ and so every $g \in \varphi(A)$ is a power of x, contradicting that G is non-cyclic and proving the claim.

Fixing such $g, h \in \varphi(A)$ satisfying (4.3), there exist two letters $a, b \in A$ with $g = \varphi(a)$ and $h = \varphi(b)$. Denoting the orders of g and h by $|g|$ and $|h|$ respectively, the set

$$\{(ab^{i|h|}a^{|g|-1}, \varepsilon_B) \in A^* \times B^* : i \in \mathbb{N}\} \tag{4.4}$$

is contained in $\Pi(\varphi, \psi)$, as

$$\varphi(ab^{i|h|}a^{|g|-1}) = gh^{i|h|}g^{|g|-1} = g(h^{|h|})^i g^{|g|-1} = gg^{|g|-1} = 1_G = \psi(\varepsilon_B)$$

for all $i \in \mathbb{N}$. Moreover, we claim any element of (4.4) is indecomposable in $\Pi(\varphi, \psi)$. Suppose for a contradiction that

$$(ab^{i|h|}a^{|g|-1}, \varepsilon_B) = (u_1, v_1)(u_2, v_2) \dots (u_p, v_p) \tag{4.5}$$

is a non-trivial decomposition of $(ab^{i|h|}a^{|g|-1}, \varepsilon_B)$ into a product of elements $(u_j, v_j) \in A^* \times B^*$. It must be that every $v_j = \varepsilon_B$, and hence $\varphi(u_j) = 1_G$ for all j. In particular, the proper prefix u_1 of $ab^{i|h|}a^{|g|-1}$ maps to 1_G.

As $\varphi(ab^j) = gh^j \neq 1$ for all $0 \leq j \leq i|h|$ by (4.4), then $u_1 = ab^{i|h|}a^j$ for some $1 \leq j \leq |g| - 1$. However, as

$$\varphi(ab^{i|h|}a^j) = gh^{i|h|}g^j = g^{j+1},$$

then $\varphi(ab^{i|h|}a^j)$ cannot equal 1_G for any $1 \leq j < |g| - 1$ by minimality of $|g|$. Hence it must be that $u_1 = ab^{i|h|}a^{|g|-1}$, which is a contradiction as (4.5) must then be a trivial decomposition. Hence as i was arbitrary, every element of the set

$$\{(ab^{i|h|}a^{|g|-1}, \varepsilon_B) \in A^* \times B^* : i \in \mathbb{N}\}$$

is indecomposable, finishing the proof. □

We now utilise Lemma 4.1.2 and Lemma 4.1.3 to give a full characterisation for when a fiber product of two free monoids over a finite fiber is finitely generated.

Theorem 4.1.4. *For two finite alphabets A and B, let $\varphi : A^* \to M$ and $\psi : B^* \to M$ be epimorphisms onto a finite monoid M. Then $\Pi(\varphi, \psi)$ is finitely generated if and only if $|\varphi(A)| = |\psi(B)| = 1$, and M is a finite cyclic group.*

Proof. ($\Rightarrow$) We prove the contrapositive. If M is not a cyclic group, then $\Pi(\varphi, \psi)$ is not finitely generated by Lemma 4.1.2 and Lemma 4.1.3. If M is a cyclic group, then it has a group presentation

$$M = \langle x : x^n = 1_M \rangle,$$

and assume first that $|\varphi(A)| > 1$. We will proceed by constructing an infinite set of indecomposable elements of $\Pi(\varphi, \psi)$, which will be enough to show that $\Pi(\varphi, \psi)$ is not finitely generated.

As $|\varphi(A)| > 1$, we can choose $a, a' \in A$ such that $\varphi(a) \neq \varphi(a')$. We will show by induction that for all $k \in \mathbb{N}$, we can construct a length k word $u_k = a_1 a_2 \dots a_k \in \{a, a'\}^+$ such that $\varphi(a_1 \dots a_i) \neq 1_M$ for all $1 \leq i \leq k$. That is, no non-empty prefix of the word u_k maps to 1_M.

For the base case, either $\varphi(a) \neq 1_M$ or $\varphi(a') \neq 1_M$ as $\varphi(a) \neq \varphi(a')$. Without loss of generality, we can set $u_1 = a$, and $\varphi(a_1) \neq 1_M$.

Assume for the inductive hypothesis that we can find a word $u_{k-1} = a_1 a_2 \dots a_{k-1} \in \{a, a'\}^+$ of length $k-1$ with $\varphi(a_1 \dots a_i) \neq 1_M$ for all $1 \leq i \leq k-1$. If

$$\varphi(a_1 a_2 \dots a_{k-1} a) = \varphi(a_1 a_2 \dots a_{k-1} a') = 1_M,$$

then

$$\varphi(a_1 a_2 \dots a_{k-1})\varphi(a) = \varphi(a_1 a_2 \dots a_{k-1})\varphi(a'),$$

implying $\varphi(a) = \varphi(a')$, which is a contradiction. Hence either $\varphi(a_1 a_2 \dots a_{k-1} a) \neq 1_M$ or $\varphi(a_1 a_2 \dots a_{k-1} a') \neq 1_M$. Both are words of length k, and by the inductive hypothesis, $\varphi(a_1 \dots a_i) \neq 1_M$ for $1 \leq i \leq k$. Making the appropriate choice of a or a' to extend the word u_{k-1} to a length k word u_k with $\varphi(u_k) \neq 1_M$ finishes the induction.

As M is a group and φ is surjective, for a given word u_k constructed as above, there is a corresponding minimal length word $v_k \in A^+$ (not necessarily of length k) such that

$$\varphi(u_k v_k) = \varphi(u_k)\varphi(v_k) = 1_M.$$

Let $w_k = u_k v_k$. As any non-empty prefix u of w_k with length less than or equal to k is a prefix of u_k, then $\varphi(u) \neq 1_M$. Moreover, as v_k is chosen to be the minimal word such that $\varphi(u_k v_k) = 1_M$, then it follows that $\varphi(u) \neq 1_M$ for any non-empty proper prefix u of w_k. Hence the infinite set of elements

$$\{(w_k, \varepsilon_B) : k \in \mathbb{N}\} \subseteq A^* \times B^* \tag{4.6}$$

is a subset of $\Pi(\varphi, \psi)$, as $\varphi(w_k) = 1_M$ for all $k \in \mathbb{N}$. Moreover, any given (w_k, ε_B) can have no non-trivial decompositions in $\Pi(\varphi, \psi)$, as every non-empty proper prefix u of w_k has the property that $\varphi(u) \neq 1_M$.

Hence (4.6) is an infinite set of indecomposable elements of $\Pi(\varphi, \psi)$ which must be contained in any generating set, showing that $\Pi(\varphi, \psi)$ is not finitely

generated. A similar proof gives that $|\psi(B)| > 1$ implies $\Pi(\varphi, \psi)$ is not finitely generated.

($\Leftarrow$) Suppose M is a cyclic group with group presentation

$$M = \langle x : x^n = 1 \rangle,$$

and $|\varphi(A)| = |\psi(B)| = 1$. By Lemma 1.6.9, $\varphi(A), \psi(B)$ are monoid generating sets for M, and hence also group generating sets. As M is cyclic, it must follow that $\varphi(A) = \{x^p\}$ and $\psi(B) = \{x^q\}$ for some $1 \leq p, q < n$ with $\gcd(p, n) = \gcd(q, n) = 1$. A pair $(u, v) \in A^* \times B^*$ is also an element of $\Pi(\varphi, \psi)$ if

$$\begin{aligned} & \varphi(u) = \psi(v) \\ \Leftrightarrow \quad & (x^p)^{|u|} = (x^q)^{|v|} \\ \Leftrightarrow \quad & x^{p|u| \,(\mathrm{mod}\, n)} = x^{q|v| \,(\mathrm{mod}\, n)} \\ \Leftrightarrow \quad & p|u| \,(\mathrm{mod}\, n) = q|v| \,(\mathrm{mod}\, n). \end{aligned}$$

Hence

$$\Pi(\varphi, \psi) = \{(u, v) \in A^* \times B^* : p|u| \equiv q|v| \,(\mathrm{mod}\, n)\}.$$

We will show that every non-identity element of $\Pi(\varphi, \psi)$ as decribed above can be written as a finite product of elements from the set

$$\begin{aligned} X = \ & \{(u, v) \in A^* \times B^* : p|u| \equiv q|v| \,(\mathrm{mod}\, n), 0 \leq |u|, |v| \leq n\} \\ & \backslash \{(u, v) \in A^* \times B^* : |u| = |v| = n \text{ or } |u| = |v| = 0\}. \end{aligned}$$

Noting that X is a subset of $\Pi(\varphi, \psi)$ which is finite, this will be sufficient to conclude the proof that $\Pi(\varphi, \psi)$ is finitely generated as a monoid.

For a given non-identity pair $(u, v) \in \Pi(\varphi, \psi)$, we can write $|u| = k_1 n + r_1$ and $|v| = k_2 n + r_2$ for some $k_1, k_2 \in \mathbb{N}^0$ and $0 \leq r_1, r_2 < n$. Hence we can write

$$u = u' x_1 x_2 \dots x_{k_1}$$

for some $u', x_i \in A^*$ with $|u'| = r_1$, $|x_i| = n$ for $1 \leq i \leq k_1$, and similarly we

can write

$$v = v' y_1 y_2 \dots y_{k_2}$$

for some $v', y_i \in B^*$ with $|v'| = r_2$, $|y_i| = n$ for $1 \leq i \leq k_2$. The pairs (x_i, ε_B) for $1 \leq i \leq k_1$ are all elements of X, as $|x_i| = n$, $|\varepsilon_B| = 0$, and

$$p|x_i| = pn \equiv 0 \equiv q0 \equiv q|\varepsilon_B| \,(\mathrm{mod}\, n).$$

Similarly, the pairs (ε_A, y_i) for $1 \leq i \leq k_2$ are all elements of X as well. As $|u'| \equiv |u| \,(\mathrm{mod}\, n)$ and $|v'| \equiv |v| \,(\mathrm{mod}\, n)$ by definition, then

$$p|u'| \equiv p|u| \equiv q|v| \equiv q|v'| \,(\mathrm{mod}\, n). \tag{4.7}$$

Hence as $0 \leq |u'| < n$, $0 \leq |v'| < n$, then the pair (u', v') is an element of X. Hence

$$(u, v) = (u', v')(x_1, \varepsilon_B)(x_2, \varepsilon_B) \dots (x_{k_1}, \varepsilon_B)(\varepsilon_A, y_1)(\varepsilon_A, y_2) \dots (\varepsilon_A, y_{k_2})$$

is a decomposition of (u, v) into a finite product of elements of X. This concludes the proof that X is a monoid generating set for $\Pi(\varphi, \psi)$, and thus $\Pi(\varphi, \psi)$ is finitely generated as a monoid. □

We now briefly give some examples of epimorphisms that satisfy the conditions of Theorem 4.1.4 (and provide a generating set for the fiber product), and that do not satisfy the conditions (finding a specific infinite set of indecomposable elements).

Examples 4.1.5. (a) Let $\{1\}$ be the trivial cyclic group, let $A = B = \{a, b\}$, and let $\varphi : A^* \to \{1\}$ be given by $\varphi(a) = \varphi(b) = 1$, extended uniquely to a homomorphism on A^* (so that $\varphi(w) = 1$ for all $w \in A^*$). Similarly, let $\psi : B^* \to \{1\}$ be given by $\psi(a) = \psi(b) = 1$, again uniquely extended to a homomorphism on B^* (so that $\varphi(w) = 0$ for all $w \in B^*$).

Then $\{1\}$ and φ, ψ match the conditions of Theorem 4.1.4, and hence $\Pi(\varphi, \psi)$ is finitely generated. In fact, as every word in A^* is mapped to 1 by φ, and every word in B^* is also mapped to 1 by ψ, then $\Pi(\varphi, \psi) = A^* \times B^*$.

By the proof of Theorem 4.1.4 (taking $p = 1$, $q = 1$, $x = 1$), a monoid generating set for $\Pi(\varphi, \psi)$ is given by

$$\begin{aligned} X &= \{(u, v) \in A^* \times B^* : |u| \equiv |v| \bmod 1, 0 \leq |u|, |v| \leq 1\} \\ &\quad \setminus \{(u, v) \in A^* \times B^* : |u| = |v| = 1 \text{ or } |u| = |v| = 0\} \\ &= \{(u, v) \in A^* \times B^* : 0 \leq |u|, |v| \leq 1, |u| \neq |v|\} \\ &= \{(\varepsilon_A, a), (\varepsilon_A, b), (a, \varepsilon_B), (b, \varepsilon_B)\}. \end{aligned}$$

For example, $(aba^2, ba) \in \Pi(\varphi, \psi)$, and can be written as the product

$$(aba^2, ba) = (a, \varepsilon_B)(b, \varepsilon_B)(a, \varepsilon_B)^2(\varepsilon_A, b)(\varepsilon_A, a).$$

(b) Let $\mathbb{Z}_2 = \{0, 1\}$ be the cyclic group of order two, let $A = B = \{a, b\}$, and let $\varphi : A^* \to \mathbb{Z}_2$ be given by $\varphi(a) = 0 = \varphi(b)$, uniquely extended to a homomorphism on A^* (so that $\varphi(w) = 0$ for all $w \in A^*$). Let $\psi : B^* \to \mathbb{Z}_2$ be given by $\psi(a) = 0$, $\psi(b) = 1$, uniquely extended to a homomorphism on B^*. In this manner, for any $w \in B^*$, we have

$$\psi(w) = \begin{cases} 1 & \text{if } w \text{ contains an odd number of } b\text{s.} \\ 0 & \text{otherwise.} \end{cases}$$

Then $\Pi(\varphi, \psi) = \{(u, v) \in A^* \times B^* : |v|_b \equiv 0 \bmod 2\}$. By Theorem 4.1.4, as $|\varphi(B)| = 2$, then $\Pi(\varphi, \psi)$ is not finitely generated. Let

$$I = \{(\varepsilon_A, ba^m b) \in A^* \times B^* : m \in \mathbb{N}\} \subseteq \Pi(\varphi, \psi).$$

If an element $(\varepsilon_A, ba^m b) \in I$ were monoid decomposable over $\Pi(\varphi, \psi)$, then it must be that

$$(\varepsilon_A, ba^m b) = (\varepsilon_A, ba^r)(\varepsilon_A, a^s b)$$

for some $0 \leq r, s \leq m$, with $r + s = m$. This cannot happen, as $\psi(ba^r) = 1$, but $\varphi(\varepsilon_A) = 0$, so $(\varepsilon_A, ba^r) \notin \Pi(\varphi, \psi)$. Hence every element of I is indecomposable, which is an example of an infinite set of indecomposables causing $\Pi(\varphi, \psi)$ to be non-finitely generated. $\triangle$

4.2 Finite presentation for fiber products of free monoids over finite fibers

Theorem 4.1.4 shows that finitely generated fiber products of two free monoids over a finite monoid are restrictive in their nature, as the epimorphisms must map every generating letter of each free monoid onto one element of a cyclic group. Given this restrictive behaviour, we can ask whether or not all such finitely generated fiber products are finitely presented, and if not, give a characterisation for the finitely presented ones. In fact, it will be our main aim for this section to prove the following theorem.

Theorem 4.2.1. *For two finite alphabets A and B , let $\varphi : A^* \to M$ and $\psi : B^* \to M$ be epimorphisms onto a finite monoid M. Then $\Pi(\varphi, \psi)$ is finitely generated if and only if it is finitely presented.*

As we wish to concern ourselves with finitely generated fiber products of free monoids over finite fibers, we will adopt the following notation as a result of Theorem 4.1.4.

Notation 4.2.2. For the rest of this section, F will be the finite cyclic group of order n, with group presentation $F = \langle x : x^n = 1 \rangle$.

A and B will be finite alphabets, ε_A and ε_B will denote the empty words in A^* and B^* respectively. The maps $\varphi : A^* \to F$ and $\psi : B^* \to F$ will be epimorphisms onto F, which satisfy $\varphi(A) = \{x^p\}$ and $\psi(B) = \{x^q\}$ for some $1 \leq p, q \leq n$ with $\gcd(p, n) = \gcd(q, n) = 1$.

$\Pi(\varphi, \psi)$ will be the fiber product of A^* and B^* with respect to φ, ψ. Note that $\Pi(\varphi, \psi)$ is finitely generated by Theorem 4.1.4.

Finally, $\bar{\Gamma}$ and Γ will be the sets of formal symbols defined

$$\bar{\Gamma} := \{\gamma(u, v) : u \in A^*, v \in B^*,\ p|u| \equiv q|v| \bmod n,\ 0 \leq |u|, |v| \leq n\}$$

and

$$\Gamma := \bar{\Gamma} \setminus \{\gamma(u, v) : u \in A^*, v \in B^*, |u| = |v| = n \text{ or } |u| = |v| = 0\}.$$

Note that if $\gamma(u,v) \in \Gamma$ with $|u| = n$, then necessarily $|v| = 0$ and hence $v = \varepsilon_B$. Conversely, if $\gamma(u,v) \in \Gamma$ with $|v| = n$, then necessarily $u = \varepsilon_A$. ■

In order to prove Theorem 4.2.1, we will first introduce two lemmas which establish a set of relations that a finitely generated fiber product of two free monoids over a finite fiber satisfies.

Lemma 4.2.3. *Recall Notation* 4.2.2. *Then the relations*

$$(\gamma(\varepsilon_A, v)\gamma(u, \varepsilon_B),\ \gamma(u, \varepsilon_B)\gamma(\varepsilon_A, v)) \quad (|u| = |v| = n); \tag{R1}$$

$$(\gamma(\varepsilon_A, v)\gamma(u, v_1),\ \gamma(u, v_2)\gamma(\varepsilon_A, v_3)) \quad (0 < |v_1| < n, |v| = |v_3| = n, \ |v_1| = |v_2|, vv_1 = v_2v_3); \tag{R2}$$

$$(\gamma(u, \varepsilon_B)\gamma(u_1, v),\ \gamma(u_2, v)\gamma(u_3, \varepsilon_B)) \quad (0 < |u_1| < n, |u| = |u_3| = n, \ |u_1| = |u_2|, uu_1 = u_2u_3) \tag{R3}$$

over Γ *hold in* $\Pi(\varphi, \psi)$.

Proof. Let $\bar{\pi} : \Gamma \to \Pi(\varphi, \psi)$ be given by $\bar{\pi}(\gamma(u,v)) = (u,v)$, and let $\pi : \Gamma^* \to \Pi(\varphi, \psi)$ be the unique homomorphism extending $\bar{\pi}$ to Γ^*, whose existence is established by Lemma 1.6.15. Then in the case of (R1), we have

$$\begin{aligned} \pi(\gamma(\varepsilon_A, v)\gamma(u, \varepsilon_B)) &= (\varepsilon_A, v)(u, \varepsilon_B) \\ &= (u, v) \\ &= (u, \varepsilon_B)(\varepsilon_A, v) \\ &= \pi(\gamma(u, \varepsilon_B)\gamma(\varepsilon_A, v)), \end{aligned}$$

and so (R1) holds in $\Pi(\varphi, \psi)$. In the case of (R2), we have

$$\begin{aligned} \pi(\gamma(\varepsilon_A, v)\gamma(u, v_1)) &= (\varepsilon_A, v)(u, v_1) \\ &= (u, vv_1) \\ &= (u, v_2v_3) \\ &= (u, v_2)(\varepsilon_A, v_3) \\ &= \pi(\gamma(u, v_2)\gamma(\varepsilon_A, v_3)), \end{aligned}$$

and hence (R2) holds in $\Pi(\varphi, \psi)$. In the case of (R3), we have

$$\begin{aligned}\pi(\gamma(u, \varepsilon_B)\gamma(u_1, v)) &= (u, \varepsilon_B)(u_1, v)\\ &= (uu_1, v)\\ &= (u_2u_3, v)\\ &= (u_2, v)(u_3, \varepsilon_B)\\ &= \pi(\gamma(u_2, v)\gamma(u_3, \varepsilon_B)),\end{aligned}$$

and hence (R3) holds in $\Pi(\varphi, \psi)$. □

Lemma 4.2.4. *Recall Notation* 4.2.2. *Let* $\gamma(u_1, v_1), \gamma(u_2, v_2) \in \Gamma$ *with* $0 < |u_1|, |u_2| < n$, $0 < |v_1|, |v_2| < n$ *and define* $u_3, u_4 \in A^*$, $v_3, v_4 \in B^*$ *in terms of* u_1, u_2, v_2, v_2 *as follows:*

$$\begin{aligned}u_1u_2 = u_3u_4, \quad |u_4| = n \quad &\text{if } |u_1u_2| > n,\\ v_1v_2 = v_3v_4, \quad |v_4| = n \quad &\text{if } |v_1v_2| > n.\end{aligned}$$

Then the relation

$$(\gamma(u_1, v_1)\gamma(u_2, v_2),\ w) \tag{R4}$$

over Γ *holds in* $\Pi(\varphi, \psi)$, *where*

$$\begin{aligned}w &= \gamma(u_1u_2, v_1v_2) && \text{if } |u_1u_2|, |v_1v_2| < n,\\ \text{or } w &= \gamma(u_3, v_1v_2)\gamma(u_4, \varepsilon_B) && \text{if } |u_1u_2| > n, |v_1v_2| < n,\\ \text{or } w &= \gamma(u_1u_2, v_3)\gamma(\varepsilon_A, v_4) && \text{if } |u_1u_2| < n, |v_1v_2| > n,\\ \text{or } w &= \gamma(u_3, v_3)\gamma(u_4, \varepsilon_B)\gamma(\varepsilon_A, v_4) && \text{if } |u_1u_2|, |v_1v_2| > n,\\ \text{or } w &= \gamma(u_1u_2, \varepsilon_B)\gamma(\varepsilon_A, v_1v_2) && \text{if } |u_1u_2| = |v_1v_2| = n.\end{aligned}$$

Proof. Let $\bar{\pi} : \Gamma \to \Pi(\varphi, \psi)$ be given by $\bar{\pi}(\gamma(u, v)) = (u, v)$, and let $\pi : \Gamma^* \to \Pi(\varphi, \psi)$ be the unique homomorphism extending $\bar{\pi}$ to Γ^*, whose existence is established by Lemma 1.6.15.

Let $\gamma(u_1, v_1), \gamma(u_2, v_2) \in \Gamma$ be two formal symbols with $0 < |u_1|, |u_2| < n$ and

$0 < |v_1|, |v_2| < n$. As $p|u_i| \equiv q|v_i| \,(\mathrm{mod}\, n)$ for $i = 1, 2$, then

$$p|u_1u_2| \equiv p(|u_1| + |u_2|) \equiv q(|v_1| + |v_2|) \equiv q|v_1v_2| \,(\mathrm{mod}\, n).$$

Hence if $|u_1u_2| = n$ then $q|v_1v_2| \equiv 0 \,(\mathrm{mod}\, n)$, implying $|v_1v_2| \equiv 0 \,(\mathrm{mod}\, n)$. It must then be that $|v_1v_2| = n$ also, accounting for the possible sizes of v_1 and v_2. Similarly, $|v_1v_2| = n$ implies $|u_1u_2| = n$ also. Hence there are five cases to check for $|u_1u_2|$ and $|v_1v_2|$:

Case 1: If $|u_1u_1|, |v_1v_2| < n$, then

$$\pi(\gamma(u_1, v_1)\gamma(u_2, v_2)) = (u_1, v_1)(u_2, v_2) = (u_1u_2, v_1v_2) = \pi(\gamma(u_1u_2, v_1v_2)),$$

and so (R4) holds.

Case 2: If $|u_1u_2| > n$, $|v_1v_2| < n$, then

$$\begin{aligned}
\pi(\gamma(u_1, v_1)\gamma(u_2, v_2)) &= (u_1, v_1)(u_2, v_2) \\
&= (u_1u_2, v_1v_2) \\
&= (u_3u_4, v_1v_2) \\
&= (u_3, v_1v_2)(u_4, \varepsilon_B) \\
&= \pi(\gamma(u_3, v_1v_2)\gamma(u_4, \varepsilon_B)),
\end{aligned}$$

and so (R4) holds.

Case 3: If $|u_1u_2| < n, |v_1v_2| > n$, then

$$\begin{aligned}
\pi(\gamma(u_1, v_1)\gamma(u_2, v_2)) &= (u_1, v_1)(u_2, v_2) \\
&= (u_1u_2, v_1v_2) \\
&= (u_1u_2, v_3v_4) \\
&= (u_1u_2, v_3)(\varepsilon_A, v_4) \\
&= \pi(\gamma(u_1u_2, v_3)\gamma(\varepsilon_A, v_4)),
\end{aligned}$$

and so (R4) holds.

Case 4: If $|u_1u_2|, |v_1v_2| > n$, then

$$\begin{aligned}\pi(\gamma(u_1, v_1)\gamma(u_2, v_2)) &= (u_1, v_1)(u_2, v_2)\\ &= (u_1u_2, v_1v_2)\\ &= (u_3u_4, v_3v_4)\\ &= (u_3, v_3)(u_4, \varepsilon_B)(\varepsilon_A, v_4)\\ &= \pi(\gamma(u_3, v_3)\gamma(u_4, \varepsilon_B)\gamma(\varepsilon_A, v_4)),\end{aligned}$$

and so (R4) holds.

Case 5: Finally if $|u_1u_2| = |v_1v_2| = n$, then

$$\begin{aligned}\pi(\gamma(u_1, v_1)\gamma(u_2, v_2)) &= (u_1, v_1)(u_2, v_2)\\ &= (u_1u_2, v_1v_2)\\ &= (u_1u_2, \varepsilon_B)(\varepsilon_A, v_1v_2)\\ &= \pi(\gamma(u_1u_2, \varepsilon_B)\gamma(\varepsilon_A, v_1v_2)),\end{aligned}$$

and so (R4) holds.

As this covers all possible cases, then (R4) holds in $\Pi(\varphi, \psi)$ as claimed. □

We now use Lemma 4.2.3 and Lemma 4.2.4 to prove Theorem 4.2.1, which we restate below.

Theorem 4.2.1. *For two finite alphabets A and B , let $\varphi : A^* \to M$ and $\psi : B^* \to M$ be epimorphisms onto a finite monoid M. Then $\Pi(\varphi, \psi)$ is finitely generated if and only if it is finitely presented.*

Proof. ($\Leftarrow$) is immediate from the definition of finite presentation.

($\Rightarrow$) Recall Notation 4.2.2, let $\bar{\pi} : \Gamma \to \Pi(\varphi, \psi)$ be given by $\bar{\pi}(\gamma(u, v)) = (u, v)$, and let $\pi : \Gamma^* \to \Pi(\varphi, \psi)$ be the unique homomorphism extending $\bar{\pi}$ to Γ^*, whose existence is established by Lemma 1.6.15.

For the set R of relations (R1)-(R4), we claim that $\Pi(\varphi, \psi)$ has monoid

presentation given by

$$\Pi(\varphi, \psi) \cong \langle \Gamma : R \rangle. \tag{4.8}$$

As R and Γ are finite, this will be enough to prove the Theorem. As $\pi(\Gamma)$ is a monoid generating set for $\Pi(\varphi, \psi)$ by the proof of Theorem 4.1.4, then $\pi(\Gamma^*) = \Pi(\varphi, \psi)$, and hence by the first isomorphism theorem

$$\Gamma^* / \ker \pi \cong \Pi(\varphi, \psi).$$

As $\langle \Gamma : R \rangle = \Gamma^* / R^\sharp$ by definition (where $R^\sharp$ is the smallest congruence on Γ^* containing R), we will show that $\ker \pi = R^\sharp$, which will suffice to prove the claim in (4.8).

As the set of relations R over Γ hold in $\Pi(\varphi, \psi)$ by Lemma 4.2.3 and Lemma 4.2.4, then $R \subseteq \ker \pi$. Hence as $\ker \pi$ is a congruence containing R, it follows that $R^\sharp \subseteq \ker \pi$.

To show that $\ker \pi \subseteq R^\sharp$, we first make the following two claims:

Claim 1: For all $w \in \Gamma^*$, $(w, w_1 w_2 w_3) \in R^\sharp$ for some

$$\begin{aligned} w_1 &\in \{\gamma(u, v) \in \Gamma : 0 < |u|, |v| < n\}^*, \\ w_2 &\in \{\gamma(u, v) \in \Gamma : |u| = n\}^*, \\ w_3 &\in \{\gamma(u, v) \in \Gamma : |v| = n\}^*. \end{aligned} \tag{4.9}$$

Claim 2: If $(w, w') \in \ker \pi$, then $(w, w_1 w_2 w_3), (w', w_1' w_2' w_3')) \in R^\sharp$ for some

$$\begin{aligned} w_1, w_1' &\in \{\gamma(u, v) \in \Gamma : 0 < |u|, |v| < n\}^*, \\ w_2, w_2' &\in \{\gamma(u, v) \in \Gamma : |u| = n\}^*, \\ w_3, w_3' &\in \{\gamma(u, v) \in \Gamma : |v| = n\}^* \end{aligned} \tag{4.10}$$

with $w_1 = w_1', w_2 = w_2', w_3 = w_3'$.

Claim 1 will be used to prove Claim 2, which is sufficient to prove that $\ker \pi \subseteq R^\sharp$ by symmetry and transitivity of $R^\sharp$, and hence finish the proof of the theorem. It thus remains to prove these two claims, which we now do.

Proof of Claim 1: For the purposes of this proof, we will refer to those elements $\gamma(u,v) \in \Gamma$ with either $u = \varepsilon_A$ or $v = \varepsilon_B$ as ε-type letters of Γ. Otherwise, $\gamma(u,v) \in \Gamma$ will be referred to as ϕ-type.

Let $w \in \Gamma^*$. We will use (R1)-(R4) in the following rewriting procedure on w (described by (A1) to (A3)) to prove the claim. An example of this rewriting procedure is given immediately following this theorem in Example 4.2.5.

(A1) If w contains a letter $\gamma(u_1,v_1)$ of ϕ-type which is immediately preceded by a letter $\gamma(u_2,v_2)$ of ε-type, then by (R2) and (R3) there exists a letter $\gamma(u_3,v_3)$ of ϕ-type, and a letter $\gamma(u_4,v_4)$ of ε-type such that

$$(\gamma(u_2,v_2)\gamma(u_1,v_1), \gamma(u_3,v_3)\gamma(u_4,v_4)) \in R.$$

Repeatedly using a sequence of (R2) and (R3) on all such ϕ-type letters which are immediately preceded by ε-type letters allows us to rewrite w as $w_{(1)}w_{(2)}$, where $w_{(1)}$ is a word consisting of ϕ-type letters (which is possibly empty if w contains no ϕ-type letters), and $w_{(2)}$ is a word consisting of ε-type letters (which again is possibly empty if w does not contain letters of ε-type). Hence

$$(w, w_{(1)}w_{(2)}) \in R^\sharp \tag{4.11}$$

for some $w_{(1)} \in \{\gamma(u,v) \in \Gamma : 0 < |u|,|v| < n\}^*$, and for some $w_{(2)} \in \{\gamma(u,v) \in \Gamma : u = \varepsilon_A \text{ or } v = \varepsilon_B\}^*$.

(A2) If $|w_{(1)}| \geq 2$, then $w_{(1)}$ contains a ϕ-type letter $\gamma(u_1,v_1)$ which is immediately followed by another ϕ-type letter $\gamma(u_2,v_2)$. Then by (R4), there is a word $w_{(3)}$ starting with a single ϕ-type letter, immediately followed by at most two ε-type letters such that

$$(\gamma(u_1,v_1)\gamma(u_2,v_2), w_{(3)}) \in R.$$

Repeatedly using (R4) from right to left on the letters of $w_{(1)}$ allows us to rewrite $w_{(1)}w_{(2)}$ as $w_1w_{(4)}$ for w_1 a ϕ-type letter (or ε_Γ), and $w_{(4)}$ a (potentially

empty) word consisting of ε-type letters. Hence

$$(w_{(1)}w_{(2)}, w_1w_{(4)}) \in R^\sharp \tag{4.12}$$

for some $w_1 \in \{\gamma(u,v) \in \Gamma : 0 < |u|, |v| < n\} \cup \{\varepsilon_\Gamma\}$, and some $w_{(4)} \in \{\gamma(u,v) \in \Gamma : u = \varepsilon_A \text{ or } v = \varepsilon_B\}^*$.

(A3) The relation (R1) describes commutativity between the letters $\gamma(u_1, v_1)$ of ε-type with $u_1 = \varepsilon_A$, and letters $\gamma(u_2, v_2)$ of ε-type with $v_2 = \varepsilon_B$. As $w_{(4)}$ is a product of letters of ε-type, then we can repeatedly use (R1) on the letters of $w_{(4)}$.

This allows us to rewrite $w_{(4)}$ as w_2w_3, where w_2 is a (possibly empty) word consisting of ε-type letters $\gamma(u,v)$ with $v = \varepsilon_B$, and w_3 is a (possibly empty) word consisting of ε-type letters $\gamma(u,v)$ with $u = \varepsilon_A$. Hence

$$(w_1w_{(4)}, w_1w_2w_3) \in R^\sharp, \tag{4.13}$$

where $w_2 \in \{\gamma(u,v) \in \Gamma : |u| = n\}^*$, and $w_3 \in \{\gamma(u,v) \in \Gamma : |v| = n\}^*$.

At the end of this rewriting procedure, combining (4.11), (4.12) and (4.13) by transitivity of $R^\sharp$ finishes the proof of Claim 1.

Proof of Claim 2: Let $(w, w') \in \ker \pi$. The fact that

$$(w, w_1w_2w_3), (w', w_1'w_2'w_3') \in R^\sharp$$

for w_i, w_i' as described in (4.10) follows from Claim 1. It remains to show that $w_i = w_i'$ for $i = 1, 2, 3$. We will do this by considering the lengths of the first and second coordinates of $\pi(w)$ and $\pi(w')$ respectively.

Let $\pi_{A^*} : \Pi(\varphi, \psi) \to A^*$ and $\pi_{B^*} : \Pi(\varphi, \psi) \to B^*$ be projections onto the first and second coordinates respectively. As

$$\pi(w) = \pi(w_1)\pi(w_2)\pi(w_3) = \pi(w_1')\pi(w_2')\pi(w_3') = \pi(w'),$$

then by considering the lengths of the first coordinates of both $\pi(w)$ and

$\pi(w')$, we have

$$|\pi_{A^*}(\pi(w))| = |\pi_{A^*}(\pi(w_1)\pi(w_2)\pi(w_3))| = |\pi_{A^*}(\pi(w_1))| + |\pi_{A^*}(\pi(w_2))|$$

as $|\pi_{A^*}(\pi(w_3))| = |\varepsilon_A| = 0$, and similarly

$$|\pi_{A^*}(\pi(w'))| = |\pi_{A^*}(\pi(w'_1)\pi(w'_2)\pi(w'_3))| = |\pi_{A^*}(\pi(w'_1))| + |\pi_{A^*}(\pi(w'_2))|.$$

As $\pi(w) = \pi(w')$, then it follows that

$$|\pi_{A^*}(\pi(w_1))| + |\pi_{A^*}(\pi(w_2))| = |\pi_{A^*}(\pi(w'_1))| + |\pi_{A^*}(\pi(w'_2))|.$$

As $w_2, w'_2 \in \{\gamma(u,v) \in \Gamma : |u| = n\}^*$, then $|\pi_{A^*}(\pi(w_2))| = n|w_2|$ and $|\pi_{A^*}(\pi(w'_2))| = n|w'_2|$, hence

$$|\pi_{A^*}(\pi(w_1))| + n|w_2| = |\pi_{A^*}(\pi(w'_1))| + n|w'_2|. \tag{4.14}$$

Similarly, by considering the lengths of the second coordinates of both $\pi(w)$ and $\pi(w')$, we see that

$$|\pi_{B^*}(\pi(w_1))| + n|w_3| = |\pi_{B^*}(\pi(w'_1))| + n|w'_3|. \tag{4.15}$$

We first consider the case of w_1 and w'_1. If $w_1 = \varepsilon_\Gamma$, then necessarily $\pi(w_1) = (\varepsilon_A, \varepsilon_B)$, and hence $|\pi_{A^*}(\pi(w_1))| = 0$. Now by taking congruences modulo n on (4.14), we see that

$$|\pi_{A^*}(\pi(w'_1))| \equiv 0 \bmod n,$$

and as $w'_1 \in \{\gamma(u,v) \in \Gamma : 0 < |u|, |v| < n\}^*$, it must be that $w'_1 = \varepsilon_{\Gamma^*}$. A similar proof shows that if $w'_1 = \varepsilon_\Gamma$, then $w_1 = \varepsilon_\Gamma$. Hence $w_1 = w'_1$ if either is equal to ε_Γ. Otherwise, it must be that $w_1 = \gamma(u,v)$ and $w'_1 = \gamma(u',v')$ for some $0 < |u|, |v|, |u'|, |v'| < n$.

Again considering (4.14) modulo n, we see that $|u| = |u'| \bmod n$, which implies that $|u| = |u'|$ on account of their sizes. As u and u' are both length $|u|$ prefixes of $\pi_{A^*}(\pi(w))$ and $\pi_{A^*}(\pi(w'))$ respectively, and $\pi_{A^*}(\pi(w)) = \pi_{A^*}(\pi(w'))$, it must be that $u = u'$. A similar proof taking

congruences modulo n on (4.15) shows that $v = v'$, and hence $w_1 = w_1'$.

In the case of w_2 and w_2', as $w_2, w_2' \in \{\gamma(u, v) \in \Gamma : |u| = n\}^*$, then $\pi(w_2) = (u, \varepsilon_B)$ and $\pi(w_2') = (u', \varepsilon_B)$ for some $u, u' \in A^*$ with $|u| \equiv |u'| \equiv 0 \bmod n$. By (4.14), as $w_1 = w_1'$, then $|w_2| = |w_2'|$. Now as u and u' are both length $n|w_2|$ suffixes of $\pi_{A^*}(\pi(w))$ and $\pi_{A^*}(\pi(w'))$ respectively, and $\pi_{A^*}(\pi(w)) = \pi_{A^*}(\pi(w'))$, it must be that $u = u'$. Hence $w_2 = w_2'$.

A very similar proof shows that $w_3 = w_3'$ when considering the second coordinates rather than the first. This concludes the proof of Claim 2, and hence of the theorem. □

We finish the section with an example of a presentation for a finitely generated fiber product of two free monoids over a finite fiber, which includes an example of the rewriting procedure described in the proof of Theorem 4.2.1.

Example 4.2.5. Let $\{1\}$ be the trivial group, let $A = B = \{a, b\}$, let $\varphi : A^* \to \{1\}$ be the constant mapping $\varphi(w) = 1$ for all $w \in A^*$, and let $\psi : B^* \to \{1\}$ be the constant mapping $\psi(w) = 1$ for all $w \in B^*$.

We saw in Examples 4.1.5 (a) that $\Pi(\varphi, \psi)$ is equal to $A^* \times B^*$, and was finitely generated by the set

$$X = \{(\varepsilon_A, a), (\varepsilon_A, b), (a, \varepsilon_B), (b, \varepsilon_B)\}.$$

We will find a finite monoid presentation for $A^* \times B^*$. As in the proof of Theorem 4.2.1, the set of formal symbols Γ will be given by

$$\Gamma = \{\gamma(\varepsilon_A, a), \gamma(\varepsilon_A, b), \gamma(a, \varepsilon_B), \gamma(b, \varepsilon_B)\}.$$

Let $x = \gamma(\varepsilon_A, a)$, $y = \gamma(\varepsilon_A, b)$, $z = \gamma(a, \varepsilon_B)$ and $t = \gamma(b, \varepsilon_B)$. Then considering Lemma 4.2.3 and Lemma 4.2.4, the relations on Γ of the form (R1) are given by the set

$$\{(xz, zx), (xt, tx), (yz, zy), (yt, ty)\},$$

and as $n = 1$, there are no relations of the form (R2), (R3) or (R4). Hence $R = \{(xz, zx), (xt, tx), (yz, zy), (yt, ty)\}$, and $\Pi(\varphi, \psi) = A^* \times B^*$ has presentation

$$\Pi(\varphi, \psi) \cong \langle x, y, z, t : (xz, zx), (xt, tx), (yz, zy), (yt, ty) \rangle$$

by the proof of Theorem 4.2.1.

We now using the writing procedure (A1)-(A4) described in Theorem 4.2.1 on the word $w = xzyt \in \Gamma^*$. Firstly, as there are no relations of the form (R2) or (R3) on Γ, then "using a sequence of (R2) and (R3) on all such ϕ-type letters" as described in (A1) is redundant. Hence w is unchanged by (A1). Similarly, as there are no relations of the form (R4), then "Repeatedly using (R4) from right to left on the letter of $w_{(1)}$" as instructed in (A2) is also redundant. Hence w is unchanged by (A2).

As in (A4), however, we can repeatedly use (R4) on w to rewrite w as $w_2 w_3$, where

$$w_2 \in \{\gamma(u, v) \in \Gamma : |u| = n\}^* = \{z, t\}^*$$

and

$$w_3 \in \{\gamma(u, v) \in \Gamma : |v| = n\}^* = \{x, y\}^*.$$

We do this in the following steps.

(i) As (xz, zx) is in R (by (R4)) and hence also in $R^\sharp$, and $R^\sharp$ is a congruence, then

$$(xz, zx)(yt, yt) = (xzyt, zxyt) \in R^\sharp. \tag{4.16}$$

(ii) As (yt, ty) is in R (by (R4)) and hence in $R^\sharp$, then

$$(zx, zx)(yt, ty) = (zxyt, zxty) \in R^\sharp. \tag{4.17}$$

(iii) As (xt, tx) is in R (by (R4)) and hence in $R^\sharp$, then

$$(z, z)(xt, tx)(y, y) = (zxty, ztxy) \in R^\sharp. \tag{4.18}$$

(iv) Finally, combining (4.16), (4.17) and (4.18) by transitivity of $R^\sharp$, we

have $(xzyt, ztxy) \in R^\sharp$.

We have now reached the end of the rewriting procedure, and have rewritten $xzyt$ as $ztxy$. $\triangle$

4.3 Finite generation for fiber products of free semigroups over infinite fibers

In Section 4.1, we classified the epimorphisms and finite fibers which result in a finitely generated fiber product of two free monoids, and further showed in Section 4.2 that all such fiber products are also finitely presented.

In this section, we move from monoids to semigroups, and it is for the same reasons highlighted in Section 4.1 that we again focus on finite generation and presentation for fiber products of free semigroups. As a consequence of Proposition 4.1.1, we will again consider the free semigroups A^+ and B^+ when A and B are finite alphabets.

We wish to ask the same questions of fiber products of free semigroups over finite fibers that we did for free monoids in Section 4.1. However, we now show that fiber products of free semigroups over a finite fiber are not finitely generated in the following result.

Proposition 4.3.1. *For two finite alphabets A and B, let $\varphi : A^+ \to S$ and $\psi : B^+ \to S$ be epimorphisms onto a finite semigroup S. Then $\Pi(\varphi, \psi)$ is not finitely generated, and hence is also not finitely presented.*

Proof. We will find an infinite subset of $\Pi(\varphi, \psi)$ consisting of indecomposable elements, which must hence be contained in any generating set. Fix any $(u, v) \in \Pi(\varphi, \psi)$. Then $\varphi(u) = \psi(v) = s$ for some $s \in S$.

As S is a finite semigroup, then s has some idempotent power s^k for $k \in \mathbb{N}$ by Lemma 1.2.10. As

$$\varphi(u^{ik}) = \varphi(u)^{ik} = (s^k)^i = s^k = \psi(v)^k = \psi(v^k),$$

for all $i \in \mathbb{N}$, then the set

$$\{(u^{ik}, v^k) : i \in \mathbb{N}\} \tag{4.19}$$

is a subset of $\Pi(\varphi, \psi)$. Suppose for a contradiction that there were a finite generating set

$$X = \{(u_i, v_i) : 1 \leq i \leq p\} \subseteq A^+ \times B^+$$

for $\Pi(\varphi, \psi)$. Then any element (u^{ik}, v^k) from the set in (4.19) can be written as a finite product of elements from X.

As v^k can be decomposed in B^+ into a product of at most $k|v|$ elements of B^+, then the element (u^{ik}, v^k) can be decomposed into a product of at most $k|v|$ elements of X. As

$$(u^{ik}, v^k) = (u_{i_1}, v_{i_1})(u_{i_2}, v_{i_2}) \dots (u_{i_q}, v_{i_q})$$

for some elements $(u_{i_j}, v_{i_j}) \in X$ with $q \leq |v|$, then it follows that

$$|u^{ik}| \leq lk|v|,$$

where $l = \max_{1 \leq i \leq p} |u_i|$. This is a contradiction, as the above inequality must hold for all $i \in \mathbb{N}$. Hence $\Pi(\varphi, \psi)$ cannot be finitely generated. □

Proposition 4.3.1 gives a stark contrast to the case of finite generation and presentation for free monoids given in the results of Theorem 4.1.4 and Theorem 4.2.1. We obtain the following necessary condition for finite generation of a fiber product of two free semigroups as a direct corollary to Proposition 4.3.1.

Corollary 4.3.2. *For two finite alphabets A and B, let $\varphi : A^+ \to S$ and $\psi : B^+ \to S$ be epimorphisms onto a semigroup fiber S. If $\Pi(\varphi, \psi)$ is finitely generated, then S is infinite.* □

We now wish to obtain further necessary conditions on a semigroup fiber S for a fiber product of two free semigroups over S to be finitely generated. We begin by showing that any fiber S for a finitely generated fiber product

of two free semigroups cannot contain idempotents.

Proposition 4.3.3. *For two finite alphabets A and B, let $\varphi : A^+ \to S$ and $\psi : B^+ \to S$ be epimorphisms onto a semigroup fiber S. If $\Pi(\varphi, \psi)$ is finitely generated, then S is idempotent-free.*

Proof. We prove the contrapositive, hence assume that S has an idempotent $e \in S$. In particular, $e^i = e$ for all $i \in \mathbb{N}$. As φ, ψ are surjections, then there exists $u \in A^+$, $v \in B^+$ such that $\varphi(u) = e = \psi(v)$. As

$$\varphi(u^i) = \varphi(u)^i = e^i = e = \psi(v)$$

for all $i \in \mathbb{N}$, it follows that the set

$$\{(u^i, v) : i \in \mathbb{N}\}$$

is a subset of $\Pi(\varphi, \psi)$. Suppose for a contradiction that $\Pi(\varphi, \psi)$ were finitely generated, and let

$$X = \{(u_i, v_i) : 1 \leq i \leq p\} \subset A^+ \times B^+$$

be any finite generating set for $\Pi(\varphi, \psi)$. Then for any $i \in \mathbb{N}$, the element (u^i, v) is expressible as a finite product of elements of X.

As v is decomposable into at most $|v|$ factors in B^+, it follows that (u^i, v) is decomposable into a product of at most $|v|$ elements of X. As

$$(u^i, v) = (u_{i_1}, v_{i_1})(u_{i_2}, v_{i_2}) \dots (u_{i_q}, v_{i_q})$$

for some elements $(u_{i_j}, v_{i_j}) \in X$ with $q \leq k|v|$, then it follows that

$$|u^i| \leq l|v|,$$

where $l = \max_{1 \leq i \leq p} |u_i|$. But this is a contradiction, as the above inequality must hold for all $i \in \mathbb{N}$. Hence $\Pi(\varphi, \psi)$ is not finitely generated. □

Proposition 4.3.3 in particular implies that S cannot be a group, and hence investigating Green's relations on the associated fiber S (which we defined in

Section 1.4) is our next aim. In particular, we will show that Green's relation $\mathcal{J}$ on S must be trivial (recalling Definition 1.4.2) for $\Pi(\varphi, \psi)$ to be finitely generated in the following result.

Proposition 4.3.4. *For two finite alphabets A and B, let $\varphi : A^+ \to S$ and $\psi : B^+ \to S$ be epimorphisms onto a semigroup fiber S. If $\Pi(\varphi, \psi)$ is finitely generated, then S is $\mathcal{J}$-trivial.*

Proof. We will prove the contrapositive, hence suppose that S is not $\mathcal{J}$-trivial. Then there exists $s, t \in S$ with $s \neq t$, but $(s, t) \in \mathcal{J}$. By Lemma 1.4.4, there exist $x, x', y, y' \in S^1$ with

$$s = xty \text{ and } t = x'sy'. \tag{4.20}$$

In particular, it follows that

$$\begin{aligned} s &= xty \\ &= (xx')s(y'y) \\ &= (xx')^2 s(y'y)^2 \\ &\vdots \\ &= (xx')^i s(y'y)^i \text{ for all } i \in \mathbb{N}. \end{aligned} \tag{4.21}$$

If S is a monoid, then in particular it has an idempotent and hence $\Pi(\varphi, \psi)$ is not finitely generated by Proposition 4.3.3.

If S is not a monoid, then $S^1 = S \cup \{1\}$. As $s \neq t$, it follows from (4.20) that at most one of x and y can equal 1, and similarly at most one of x' and y' can equal 1.

As $mn = 1 \Leftrightarrow m = 1$ and $n = 1$ for all $m, n \in S \cup \{1\}$, to avoid a contradiction it must be that at most one of xx' and $y'y$ can equal 1. We consider all of the possible cases:

Case 1: $xx' \neq 1$ and $y'y \neq 1$. In this case, it follows that both $xx' \in S$ and $y'y \in S$.

As φ, ψ are surjections, then there exist $u, w, w' \in A^+$, $v \in B^+$ with $\varphi(u) = s = \psi(v)$ and $\varphi(w) = xx'$, $\varphi(w') = y'y$. Hence by (4.21), as

$$\varphi(w^i u (w')^i) = \varphi(w)^i \varphi(u) \varphi(w')^i = (xx')^i s (y'y)^i = s = \psi(v)$$

for all $i \in \mathbb{N}$, then the set

$$\{(w^i u (w')^i, v) : i \in \mathbb{N}\}$$

is a subset of $\Pi(\varphi, \psi)$. Suppose for a contradiction that $\Pi(\varphi, \psi)$ were finitely generated, and let

$$X = \{(u_i, v_i) : 1 \leq i \leq p\} \subset A^+ \times B^+$$

be any finite generating set. For any $i \in \mathbb{N}$, the element $(w^i u (w')^i, v)$ can be decomposed as a finite product of elements of X, and in particular must be decomposable into a product of at most $|v|$ elements of X (as v can be decomposed into a product of at most $|v|$ elements in B^+).

As

$$(w^i u (w')^i, v) = (u_{i_1}, v_{i_1})(u_{i_2}, v_{i_2}) \dots (u_{i_q}, v_{i_q})$$

for some elements $(u_{i_j}, v_{i_j}) \in X$ with $q \leq |v|$, it follows that

$$|w^i u (w')^i| \leq l|v|,$$

where $l = \max_{1 \leq i \leq p} |u_i|$. The above inequality must hold for all $i \in \mathbb{N}$ however, which leads to a contradiction. Hence it follows that $\Pi(\varphi, \psi)$ is not finitely generated in this case.

Case 2: $xx' = 1$. In this case, $y'y \neq 1$, or equivalently $y'y \in S$. As

$$s = s(y'y)^i$$

for all $i \in \mathbb{N}$ by (4.20), then the set

$$\{(u(w')^i, v) : i \in \mathbb{N}\}$$

(where u, w', v are as in Case 1 is a subset of $\Pi(\varphi, \psi)$, similarly to Case 1. The proof by contradiction that $\Pi(\varphi, \psi)$ is not finitely generated is now similar to Case 1, ignoring instances of the word w.

Case 3: $y'y = 1$. In this case, $xx' \neq 1$, or equivalently $xx' \in S$. As

$$s = (xx')^i s$$

for all $i \in \mathbb{N}$ by (4.20), then the set

$$\{(w^i u, v) : i \in \mathbb{N}\}$$

(where u, w, v are as in Case 1) is a subset of $\Pi(\varphi, \psi)$, again similarly to Case 1. Once more, the proof by contradiction that $\Pi(\varphi, \psi)$ is not finitely generated is now similar to Case 1, this time ignoring instances of the word w'.

This concludes all of the cases, and hence in any situation, $\Pi(\varphi, \psi)$ is not finitely generated. This completes the proof of the contrapositive, and hence of the proposition. □

Noting that Green's relation $\mathcal{J}$ contains $\mathcal{L}, \mathcal{R}$, and $\mathcal{H}$, it then follows as a corollary to the above that S must be $\mathcal{K}$-trivial for a fiber product of two free semigroups over S to be finitely generated, for $\mathcal{K}$ any Green's relation. This completely characterises Green's relations on S, and rules out classes such as groups, inverse semigroups, bands, semilattices, left (right) groups, and others (definitions of which can all be found in [16]).

Perhaps the most natural examples of semigroups which are infinite but finitely generated, idempotent-free and $\mathcal{J}$-trivial (thus satisfying the conditions of the results of this section) are the finitely generated free semigroups themselves. These will be the topic of discussion in Chapter 5.

Finitely generated free commutative semigroups of course also have these properties. We will work for the remainder of the section to prove the following result:

Theorem 4.3.5. *For two finite alphabets A and B, let $\varphi : A^+ \to S$ and $\psi : B^+ \to S$ be epimorphisms onto the free commutative semigroup S with semigroup presentation*

$$S = \langle x_1, x_2, \ldots, x_n : x_i x_j = x_j x_i, 1 \leq i, j \leq n \rangle.$$

Then $\Pi(\varphi, \psi)$ is finitely generated if and only if $n = 1$ (so that $S \cong \mathbb{N}$), and either $|\varphi(A)| = 1$ or $|\psi(B)| = 1$.

In particular, this result will demonstrate that though the required conditions found in this section are rather restrictive, they are necessary but not sufficient for finite generation through the example of free commutative semigroups. We introduce the following lemmas which will aid us in proving Theorem 4.3.5.

Lemma 4.3.6. *For two finite alphabets A and B, let $\varphi : A^+ \to S$ and $\psi : B^+ \to S$ be epimorphisms onto the free commutative semigroup S with semigroup presentation*

$$S = \langle x_1, x_2, \ldots, x_n : x_i x_j = x_j x_i, 1 \leq i, j \leq n \rangle.$$

If $n > 1$, then $\Pi(\varphi, \psi)$ is not finitely generated.

Proof. First, any semigroup generating set for S must contain $x_1, x_2, \ldots, x_n$. As $\varphi(A)$, $\psi(B)$ are semigroup generating sets for S by Lemma 1.6.10, then each of $x_1, x_2, \ldots, x_n$ are in both $\varphi(A)$ and $\psi(B)$. Thus in particular there exist $a, a' \in A$, $b, b' \in B$ with $\varphi(a) = x_1 = \psi(b)$, and $\varphi(a') = x_2 = \psi(b')$.

As

$$\varphi(a^i a') = \varphi(a)^i \varphi(a') = x_1^i x_2 = x_2 x_1^i = \psi(b')\psi(b)^i = \psi(b' b^i)$$

for all $i \in \mathbb{N}$, it follows that the set

$$\{(a^i a', b' b^i) : i \in \mathbb{N}\} \tag{4.22}$$

is a subset of $\Pi(\varphi, \psi)$. We claim that every element of (4.22) is indecomposable in $\Pi(\varphi, \psi)$, which is enough to prove that $\Pi(\varphi, \psi)$ is not finitely

generated, as any generating set must contain (4.22).

Suppose to the contrary; that an element $(a^i a', b'b^i)$ of (4.22) can be decomposed into a non-trivial product

$$(a^i a', b'b^i) = (u, v)(u', v')$$

for some $(u, v), (u', v') \in \Pi(\varphi, \psi)$. As this product is non-trivial, then u is a proper prefix of $a^i a'$, and v is a proper prefix of $b'b^i$.

The possible proper prefixes of $a^i a'$ and $b'b^i$ are of the form a^j and $b'b^k$ respectively, for some $1 \leq j \leq i$, and $0 \leq k < i$. But the images of such prefixes are

$$\varphi(a^j) = \varphi(a)^j = x_1^j,$$

and

$$\psi(b'b^k) = \psi(b')\psi(b)^k = x_2 x_1^k.$$

As $n > 1$, then $x_1 \neq x_2$, and hence $x_1^j \neq x_2 x_1^k$ for any $1 \leq j \leq i$, and $0 \leq k < i$. This is a contradiction, as $\varphi(u) \neq \psi(v)$, but $(u, v) \in \Pi(\varphi, \psi)$. Hence it must be that $\Pi(\varphi, \psi)$ is not finitely generated in the case where $n > 1$. □

We thus restrict to the case where S is the free commutative semigroup of rank one; which we will consider as the natural numbers $\mathbb{N}$ in the following lemma.

Lemma 4.3.7. *For two finite alphabets A and B, let $\varphi : A^+ \to \mathbb{N}$ and $\psi : B^+ \to \mathbb{N}$ be epimorphisms onto $\mathbb{N}$. If $|\varphi(A)|$ and $|\psi(B)| > 1$, then $\Pi(\varphi, \psi)$ is not finitely generated.*

Proof. As φ, ψ are surjections onto $\mathbb{N}$ and $\varphi(A)$ and $\psi(B)$ are semigroup generating sets for $\mathbb{N}$ by Lemma 1.6.10, then there exist some $a \in A$, $b \in B$ with $\varphi(a) = 1 = \psi(b)$. Moreover, as $|\varphi(A)|, |\psi(B)| > 1$, then there exist $a' \in A$, $b' \in B$ with $\varphi(a') = p$, $\psi(b') = q$ with $p, q > 1$. Either $p \geq q$, or $q \geq p$, which we will consider separately as cases.

Case 1: $p \geq q$. In this case, $p = cq + d$ for some $c \in \mathbb{N}$, $0 \leq d < q$. Let $w = (b')^c b^d$, hence

$$\psi(w) = \psi(b')^c \psi(b)^d = cq + d = p. \tag{4.23}$$

Now as

$$\varphi(a(a')^i) = \varphi(a)\varphi(a')^i = 1 + ip = ip + 1 = \psi(w)^i \psi(b) = \psi(w^i b),$$

then the set

$$\{(a(a')^i, w^i b) : i \in \mathbb{N}\}$$

is a subset of $\Pi(\varphi, \psi)$. We will show that any element $(a(a')^i, w^i b)$ is indecomposable in $\Pi(\varphi, \psi)$, which will be sufficient to show that $\Pi(\varphi, \psi)$ is not finitely generated (as any generating set for $\Pi(\varphi, \psi)$ must contain the set of indecomposables).

Suppose to the contrary, that $(a(a')^i, w^i b)$ were reducible. Then

$$(a(a')^i, w^i b) = (u, v)(u', v')$$

for some $(u, v), (u', v') \in \Pi(\varphi, \psi)$. Then u is a proper prefix of $a(a')^i$, hence $u = a(a')^j$ for some $0 \leq j < i$ and thus $\varphi(u) = 1 + jp$. It follows that $\varphi(u) \equiv 1 \bmod p$.

Moreover, as $\varphi(u) = \psi(v)$, then $\psi(v) \equiv 1 \bmod p$ also. As v is a proper prefix of $w^i b$, the possibilities for v are as follows:

(i) $v = w^j (b')^k$ for some $0 \leq j \leq i - 1,\ 0 \leq k < c$ with $j + k > 0$;

(ii) $v = w^j (b')^c b^l$ for some $0 \leq j \leq i - 1,\ 0 \leq l \leq d$.

In the instance of (i), we have

$$\psi(v) = \psi(w)^j \psi(b')^k = jp + kq \equiv kq \bmod p,$$

and in the instance of (ii) we have

$$\psi(v) = \psi(w)^j \psi(b')^c \psi(b)^l = jp + cq + l \equiv (cq + l) \bmod p.$$

As $kq < p$ for $0 \leq k < c$ and $cq + l \leq p$ for $0 \leq l \leq d$, then the set

$$\{kq, cq + l : 0 \leq k < c, 0 \leq l \leq d\}$$

form a set of least positive residues modulo p. As $\psi(v) \equiv 1 \bmod p$, then one of the elements of this set must equal 1. However as $q > 1$ by assumption, then there can be no $0 \leq k < c$ or $0 \leq l \leq d$ such that $kq = 1$ or $cq + l = 1$. This is a contradiction, and hence it must be that the pair $(a(a')^i, w^i b)$ is irreducible in $\Pi(\varphi, \psi)$. Now as i was arbitrary, it follows that the set

$$\{(a(a')^i, w^i b) : i \in \mathbb{N}\}$$

is an infinite set of irreducible elements of $\Pi(\varphi, \psi)$ as claimed, and hence $\Pi(\varphi, \psi)$ is not finitely generated.

Case 2: $q \geq p$. In this case, we can write $q = cp + d$ for some $c \in \mathbb{N}$, $0 \leq d < p$, and similarly to the case for $p \geq q$, it will follow that the set

$$\{(w^i a, b(b')^i) : i \in \mathbb{N}\}$$

(where $w = (a')^c a^d$) is an infinite set of irreducible elements of $\Pi(\varphi, \psi)$ by a very similar proof to Case 1; replacing all instances of $p, q, a, a', b, b', \varphi, \psi$ by $q, p, b, b', a, a', \psi, \varphi$ respectively, and reversing the coordinates of any pair. Hence again, $\Pi(\varphi, \psi)$ is not finitely generated. □

We are now ready to prove Theorem 4.3.5 as the final result of the section, which we now restate.

Theorem 4.3.5. *For two finite alphabets A and B, let $\varphi : A^+ \to S$ and $\psi : B^+ \to S$ be epimorphisms onto the free commutative semigroup S with semigroup presentation*

$$S = \langle x_1, x_2, \ldots, x_n : x_i x_j = x_j x_i, 1 \leq i, j \leq n \rangle.$$

Then $\Pi(\varphi, \psi)$ is finitely generated if and only if $n = 1$ (so that $S \cong \mathbb{N}$), and either $|\varphi(A)| = 1$ or $|\psi(B)| = 1$.

Proof. ($\Rightarrow$) We prove the contrapositive. If $n > 1$, then $\Pi(\varphi, \psi)$ is not finitely generated by Lemma 4.3.6.

Otherwise, if $n = 1$ but both $|\varphi(A)| > 1$ and $|\psi(B)| > 1$, then $\Pi(\varphi, \psi)$ is not finitely generated by Lemma 4.3.7.

($\Leftarrow$) Assume that $n = 1$, and either $|\varphi(A)| = 1$ or $|\psi(B)| = 1$. As $S \cong \mathbb{N}$, we will view the fiber as $\mathbb{N}$ and consider the epimorphisms as onto $\mathbb{N}$.

In particular, as φ, ψ are surjections and $\varphi(A)$ and $\psi(B)$ are semigroup generating sets for $\mathbb{N}$ by Lemma 1.6.10, it must be that either $\varphi(A) = \{1\}$ or $\psi(B) = \{1\}$. We will consider the case where $\varphi(A) = \{1\}$, and argue that the case for $\psi(B) = \{1\}$ is similarly proved.

In this case, $\varphi(a) = 1$ for all $a \in A$, hence $\varphi(u) = |u|$ for all $u \in A^+$. It thus follows that

$$\Pi(\varphi, \psi) = \{(u, v) \in A^+ \times B^+ : |u| = \psi(v)\}.$$

We will show that any element of $\Pi(\varphi, \psi)$ can be written as a finite product of elements of the set

$$X = \{(u, v) \in \Pi(\varphi, \psi) : v \in B\}.$$

X is finite, as there are only finitely many possibilities for $v \in B$, and any corresponding u such that $(u, v) \in X$ has bounded word length, as $|u| = \psi(v)$.

Let $(u, v) \in \Pi(\varphi, \psi)$. If $|v| = 1$, then $v \in B$ and hence $(u, v) \in X$. Otherwise, as

$$u = a_1 a_2 \dots a_{|u|},\ v = b_1 b_2 \dots b_{|v|}$$

for some $a_1, a_2, \dots, a_{|u|} \in A$ and $b_1, b_2, \dots, b_{|v|} \in B$, then for $1 \leq i \leq |v|$ let

$$w_i = a_{p_i+1} a_{p_i+2} \dots a_{p_i+\psi(b_i)}$$

where $p_1 = 0$, and $p_i = p_{i-1} + \psi(b_{i-1})$ for $2 \leq i \leq |v|$.

As $|w_i| = |a_{p_i+1}a_{p_i+2}\dots a_{p_i+\psi(b_i)}| = \psi(b_i)$, then it follows that

$$(w_i, b_i) \in X \text{ for each } 1 \leq i \leq |v|.$$

We will also argue that

$$(u, v) = (w_1, b_1)(w_2, b_2)\dots(w_{|v|}, b_{|v|}),$$

which is sufficient to prove that every element of $\Pi(\varphi, \psi)$ can be written as a finite product of elements from X, proving finite generation. As

$$(w_1, b_1)(w_2, b_2)\dots(w_{|v|}, b_{|v|}) = (w_1w_2\dots w_{|v|}, b_1b_2\dots b_{|v|}) = (w_1w_2\dots w_{|v|}, v),$$

it remains to argue that $u = w_1w_2\dots w_{|v|}$.

The first letter of w_1 is a_1. For every $1 \leq i \leq |v|$, each letter of w_i (except for the last letter) has an index which is precisely one less than that of the following letter by construction.

For all $i < |v|$, the last letter of w_i has index $p_i + \psi(b_i)$, and the first letter of w_{i+1} has index $p_{i+1} + 1$. Hence as $p_{i+1} + 1 = p_i + \psi(b_i) + 1$ by definition, the index of the last letter of w_i is precisely one less than that of the first letter of w_{i+1}.

In the case of $i = |v|$, the last letter of $w_{|v|}$ has index

$$\begin{aligned}
p_{|v|} + \psi(b_v) &= p_{|v|-1} + \psi(b_{|v|-1}) + \psi(b_{|v|}) \\
&= p_{|v|-2} + \psi(b_{|v|-2}) + \psi(b_{|v|-1}) + \psi(b_{|v|}) \\
&\vdots \\
&= \sum_{k=1}^{|v|} \psi(b_k) \\
&= \psi(b_1b_2\dots b_{|v|}) \\
&= \psi(v) \\
&= |u|
\end{aligned}$$

The above arguments on letter indices combine to give that

$$w_1 w_2 \dots w_{|v|} = a_1 a_2 \dots a_{|u|} = u,$$

which completes the proof that (u, v) can be written as a finite product of elements of X, and hence of the proof that $\Pi(\varphi, \psi)$ is finitely generated when $n = 1$ and $|\varphi(A)| = 1$. The case when $|\psi(B)| = 1$ can be proved in a symmetric manner, using the generating set

$$X = \{(u, v) \in \Pi(\varphi, \psi) : u \in A\}. \qquad \square$$

We end the section with an example of a finitely generated fiber product of two free semigroups over a free commutative semigroup

Example 4.3.8. Take $S = \mathbb{N}$. Let $A = B = \{a, b\}$, and define $\varphi : A^+ \to S$ by $\varphi(w) = |w|$ for all $w \in A^+$. Further define $\psi : B^+ \to S$ by $\psi(a) = 1$, $\psi(b) = 2$ uniquely extended to a homomorphism on B^+, so that $\psi(w) = |w|_a + 2|w|_b$ for all $w \in B^+$.

Then $\Pi(\varphi, \psi) = \{(u, v) \in A^* \times B^* : |u| = |v|_a + 2|v|_b\}$. Moreover, $\Pi(\varphi, \psi)$ is finitely generated by Theorem 4.3.5, as $S = \mathbb{N}$ and $|\varphi(A)| = 1$. As in the proof of Theorem 4.3.5, a finite generating set for $\Pi(\varphi, \psi)$ is given by

$$\begin{aligned} X &= \{(u, v) \in \Pi(\varphi, \psi) : v \in B\} \\ &= \{(a, a), (b, a), (a^2, b), (ab, b), (ba, b), (b^2, b)\} \end{aligned} \qquad \triangle$$

Chapter 5

Deciding finite generation for fiber products of free semigroups and monoids

In the previous chapter, we classified those fiber products of free semigroups/monoids over finite semigroups/monoids which are finitely generated, which were the results of Proposition 4.3.1 and Theorem 4.1.4 respectively. We saw that, in the monoid case, the structural properties of the fiber are somewhat strong whenever we have finite generation, but as a consequence implied finite presentation.

We saw in the case of free monoids, the fiber need be a cyclic group in the finite case. For finitely generated fiber products of free semigroups, we saw that the fiber need be an infinite (but finitely generated), $\mathcal{J}$-trivial idempotent free semigroup (see Proposition 4.3.1, Proposition 4.3.3, and Proposition 4.3.4).

Of course, the associated epimorphisms for the fiber product play a role too. In the cases of finite generation seen so far, we have found restrictions for at least one of the epimorphisms to be a constant map on the given alphabet (see Theorem 4.1.4, Theorem 4.3.5). Though the conditions giving finite generation for fiber products of free monoids (given in Proposition 4.3.1) are strong, the problems of finite generation and presentation for such fiber

products are *decidable* as a result.

Another notable decision problem is the membership problem (which we refer the reader to Definition 1.6.24 for recap) for a semigroup. In the scope of this work, it is a natural question to ask whether given semigroups S, T, U, epimorphisms $\varphi : S \to U$, $\psi : T \to U$, and a pair $(s, t) \in S \times T$ whether or not (s, t) belongs to the fiber product $\Pi(\varphi, \psi)$. In fact, the decidability of the word problem for the fiber is sufficient for decidability of the membership problem for a fiber product outlined, as seen in the following result.

Proposition 5.0.1. *Let S, T, U be semigroups, and let $\varphi : S \to U$ and $\psi : T \to U$ be epimorphisms onto U such that the associated fiber product $\Pi(\varphi, \psi)$ is finitely generated as a semigroup.*

If the word problem of U is decidable, then the membership problem for $\Pi(\varphi, \psi)$ in $S \times T$ is decidable.

Proof. Let X be any generating set for $S \times T$, and let w be a word over X. Letting $\pi_S : S \times T \to S$ and $\pi_T : S \times T \to T$ be projections onto S and T respectively, then $w \in \Pi(\varphi, \psi)$ if and only if $\varphi(\pi_S(w)) = \psi(\pi_T(w))$.

As $\Pi(\varphi, \psi)$ is finitely generated, then U is finitely generated by Proposition 4.1.1. Letting Y be any finite generating set for U, then we can write $\varphi(\pi_S(w))$ and $\psi(\pi_T(w))$ as words over Y. As the word problem of U is decidable, then it is decidable whether or not $\varphi(\pi_S(w))$ and $\psi(\pi_T(w))$ represent the same element of U, and hence whether or not $w \in \Pi(\varphi, \psi)$, completing the proof. □

Our results indicate that decision problems for the fiber product are linked with the properties of the fiber. Hence as the main aim for this chapter, considering the problem of finite generation, we seek an example of a fiber satisfying the conditions of Corollary 4.3.2, Proposition 4.3.3, and Proposition 4.3.4, whose fiber product has decidable finite generation problem.

Perhaps the most natural example of semigroups which are idempotent free, $\mathcal{J}$-trivial, infinite but finitely generated with decidable word problem are

the (finite rank) free semigroups themselves. Hence in this chapter, we will consider the finite generation problems for fiber products of free semigroups and monoids over free semigroup and monoid fibers.

We begin with Section 5.1, where we obtain equivalent conditions to finite generation for fiber products of free semigroups and monoids. Namely, we will show that such fiber products are finitely generated if and only if they have finitely many indecomposable elements.

Having completed this, we will then introduce the machinery of *two-tape automata* in Section 5.2, with some worked examples, and construct such automata from fiber products of free monoids over free monoid fibers. In Section 5.3, we will show the automata constructed in Section 5.2 determine the indecomposable elements of the associated fiber product.

We conclude with Section 5.4 by showing that the finite generation problem for fiber products of free monoids over free monoid fibers is decidable, using the automata. We then derive analogous results for fiber products of free semigroups over free semigroup fibers. We finish with some examples of such finitely generated fiber products.

We note that this chapter is based on Section 5 of the paper [6], written by the author.

5.1 Equivalent conditions to finite generation for fiber products of free semigroups and monoids over free fibers

As mentioned at the beginning of the chapter, we will be considering the finite generation problem for fiber products of free semigroups and monoids over free fibers. In this section, we first give the following recharacterisation of finite generation for such a fiber product of two free monoids in relation to decomposition.

Lemma 5.1.1. *For two finite alphabets A and B, let $\varphi : A^* \to M$ and*

$\psi : B^* \to M$ *be epimorphisms onto a monoid* M*. Then* $\Pi(\varphi, \psi)$ *is finitely generated as a monoid if and only if* $\Pi(\varphi, \psi)$ *has finitely many indecomposable elements.*

Proof. ($\Rightarrow$) Every generating set for $\Pi(\varphi, \psi)$ must contain the set of indecomposable elements, hence if $\Pi(\varphi, \psi)$ is finitely generated, $\Pi(\varphi, \psi)$ has finitely many indecomposable elements.

($\Leftarrow$) We will show that the set of indecomposable elements is a generating set for $\Pi(\varphi, \psi)$. Hence let $(u, v) \in \Pi(\varphi, \psi) \setminus \{(\varepsilon_A, \varepsilon_B)\}$.

If (u, v) is indecomposable, then we are done. Otherwise, if (u, v) is decomposable, we will prove by induction on the sum of the lengths of u and v that (u, v) can be written as a finite product of indecomposable elements from $\Pi(\varphi, \psi)$.

For the base case, let

$$m = \min_{\substack{(u,v)\in\Pi(\varphi,\psi),\\ (u,v)\neq(\varepsilon_A,\varepsilon_B)}} (|u| + |v|).$$

Then any pair $(u, v) \in \Pi(\varphi, \psi)$ with $|u| + |v| = m$ is necessarily indecomposable, for otherwise there is a non-trivial decomposition

$$(u, v) = (u_1, v_1)(u_2, v_2)$$

for some $(u_1, v_1), (u_2, v_2) \in \Pi(\varphi, \psi)$, but as

$$|u_1| + |v_1| < |u_1| + |u_2| + |v_1| + |v_2| = |u| + |v| = m$$

and similarly $|u_2| + |v_2| < m$, we would obtain a contradiction on the minimality of m. This proves the base case.

Assume for the inductive hypothesis that any pair $(u, v) \in \Pi(\varphi, \psi)$ with $|u| + |v| = k$ can be written as a finite product of indecomposable elements of $\Pi(\varphi, \psi)$, and consider for the next step any pair $(u', v') \in \Pi(\varphi, \psi)$ with

$|u'| + |v'| = n$, where

$$n = \min_{\substack{(u',v') \in \Pi(\varphi,\psi), \\ |u'|+|v'|>k}} (|u'| + |v'|).$$

If (u', v') is indecomposable, then there is nothing to show. Otherwise, (u', v') is non-trivially decomposable into a product

$$(u', v') = (u'_1, v'_1)(u'_2, v'_2)$$

for some $(u'_1, v'_1), (u'_2, v'_2) \in \Pi(\varphi, \psi)$. As

$$|u'_1| + |v'_1| < |u'_1| + |u'_2| + |v'_1| + |v'_2| = |u'| + |v'| = n,$$

and similarly $|u'_2| + |v'_2| < n$, it must follow from minimality of n that $|u'_1| + |v'_1| \leq k$, $|u'_2| + |v'_2| \leq k$.

By the inductive hypothesis, each of (u'_1, v'_1) and (u'_2, v'_2) can be written as a finite product of elements of $\Pi(\varphi, \psi)$, and hence so can (u', v'). This completes the proof of the induction, and hence of the lemma. □

We now give the following analogous statement for fiber products of free semigroups as a corollary to this result.

Lemma 5.1.2. *For two finite alphabets A and B, let $\varphi : A^+ \to S$ and $\psi : B^+ \to S$ be epimorphisms onto a semigroup S. Then $\Pi(\varphi, \psi)$ is finitely generated as a semigroup if and only if $\Pi(\varphi, \psi)$ has finitely many indecomposable elements.*

Proof. The proof is precisely the same as for Lemma 5.1.1, without accounting for empty words. □

Given the similarity between Lemma 5.1.1 and Lemma 5.1.2, we will concentrate on first obtaining results on finite generation for fiber products of free monoids over free fibers. We will then derive the analogous results for fiber products of free semigroups over free fibers later in the chapter.

5.2 Two-tape automata construction for fiber products of free monoids over free monoid fibers

As a consequence of Lemma 5.1.1 and Lemma 5.1.2, deciding finite generation for $\Pi(\varphi, \psi)$ is equivalent to deciding whether its set of indecomposable elements is finite. In this section, we will construct a finite automaton from a fiber product of two free monoids over a free fiber, and show in section 5.3 that it accepts a language corresponding to the fiber product's indecomposable elements. For a fiber product of two free monoids over a free monoid fiber, if the accepted language of the automaton is finite, then we will have decided that the fiber product is finitely generated as a consequence of Lemma 5.1.1.

As the fiber products under consideration are semigroups of pairs of free words, we will use *two-tape automata*, which we now define.

Definition 5.2.1. A *two-tape automaton* $\mathcal{A}$ is a 6-tuple

$$\mathcal{A} = (Q, \Sigma_1, \Sigma_2, \delta, \iota, F),$$

where:

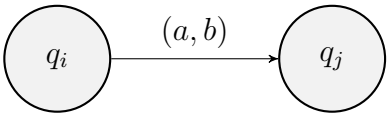
- Q is a finite set of *states*;
- Σ_1, Σ_2 are two finite alphabets forming the *input alphabet*

  $$\Sigma = (\Sigma_1 \cup \{\varepsilon_1\}) \times (\Sigma_2 \cup \{\varepsilon_2\})$$

  where $\varepsilon_1, \varepsilon_2$ are the empty words over Σ_1 and Σ_2 respectively;
- δ is a subset of $Q \times \Sigma \times Q$ called the *transition relation*;
- $\iota \in Q$ and $F \subseteq Q$ are the *initial state* and *set of final states* respectively.

Pictorially, the initial state ι and any final states $q_i \in F$ will be represented respectively by the following nodes :

All other states will be represented by a regular circular node. A transition $(q_i, (a, b), q_j)$ of δ will be represented pictorially by the *edge*



and say that the edge has *label* (a, b). Where there is more than one edge going from state q_i to state q_j, we will write each label alongside a single arrow, rather than drawing multiple arrows.

A *path* p of *length* n in $\mathcal{A}$ is a sequence of transitions $p = (q_{i-1}, \sigma_i, q_i)_{i=1}^n$. A *cycle* is a path $p = (q_{i-1}, \sigma_i, q_i)_{i=1}^n$ with $q_0 = q_n$. A two-tape automaton is said to be *acyclic* if it contains no cycles.

As $\Sigma \subseteq \Sigma_1^* \times \Sigma_2^*$, then there is a natural surjective homomorphism $\pi : \Sigma^+ \to \Sigma_1^* \times \Sigma_2^*$ defined by

$$\pi(\sigma_1\sigma_2 \ldots \sigma_n) = \sigma_1 \cdot \sigma_2 \cdot \cdots \cdot \sigma_n$$

where the right hand side is considered as a product of pairs in the *product monoid* $\Sigma_1^* \times \Sigma_2^*$. The mapping π will be called the *product monoid mapping*.

A pair $(u, v) \in \Sigma_1^* \times \Sigma_2^*$ will be a *label* for a path $p = (q_{i-1}, \sigma_i, q_i)_{i=1}^n$ if

$$\pi(\sigma_1\sigma_2 \ldots \sigma_n) = (u, v).$$

An *input* is a word $w \in \Sigma^+$. We will say that the automaton $\mathcal{A}$ *accepts* an input $w = \sigma_1\sigma_2 \ldots \sigma_n \in \Sigma^+$ if there exists a path $p = (q_{i-1}, \sigma_i, q_i)_{i=1}^n$ with $q_0 = \iota$, and $q_n \in F$. The *language accepted by* $\mathcal{A}$ is the set of all words $w \in \Sigma^+$ which $\mathcal{A}$ accepts, and will be denoted $\mathcal{L}(\mathcal{A})$. ■

Two-tape automata provide us with a way of considering pairs $(u, v) \in \Sigma_1^* \times \Sigma_2^*$ as labels of the automaton. As the input alphabet is a

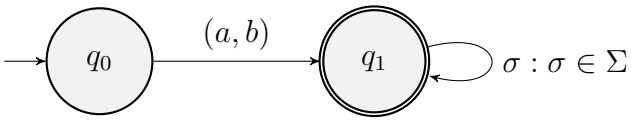
direct product, then the automaton can either read either letters from u, or from v, or from both u and v at the same time (hence being able to read from "two tapes").

We now provide an example of a two-tape automaton, illustrating the connection between the language accepted by it and the corresponding language in the product monoid.

Example 5.2.2. Let $\mathcal{A} = (Q, \Sigma_1, \Sigma_2, \delta, \iota, F)$, where

- $Q = \{q_0, q_1\}$;
- $\Sigma_1 = \Sigma_2 = \{a, b\}$;
- $\delta = \{(q_0, (a, b), q_1), (q_1, \sigma, q_1) : \sigma \in \Sigma\}$;
- $\iota = q_0, F = \{q_1\}$.

Pictorially, $\mathcal{A}$ has the following form:



We can see that every input of the form $(a, b)w$ for $w \in \Sigma^*$ is accepted by the automaton, and conversely every input accepted by the automaton has to be of this form. Hence

$$\mathcal{L}(\mathcal{A}) = (a, b)\Sigma^* = \{(a, b)w : w \in \Sigma^*\}.$$

Considering the natural mapping of this language into $\Sigma_1^* \times \Sigma_2^*$, then

$$\pi(\mathcal{L}(\mathcal{A})) = \{(u, v) \in \Sigma_1^* \times \Sigma_2^* : u_1 = a, v_1 = b\}.$$

That is, for every pair of words $(u, v) \in \Sigma_1^* \times \Sigma_2^*$ where u begins with an a, and v begins with a b, there is at least one corresponding word in Σ^* that the automaton accepts, and vice versa. △

From a given fiber product of two free monoids over a free monoid fiber, we now construct a two-tape automaton that can be used to decide whether

the fiber product is finitely generated. The automaton will be designed to accepted all input words whose image (under the map π defined in Definition 5.2.1) is an indecomposable element in the fiber product. This will be proved later in Section 5.3, but we first give the construction in the following definition, for which we ask the reader to recall the notation of Definition 1.5.3 and Definition 1.5.4.

Definition 5.2.3. Let A, B, C be finite alphabets, and let $\varphi : A^* \to C^*$, $\psi : B^* \to C^*$ be two epimorphisms. Given the fiber product $\Pi(\varphi, \psi)$, the *two-tape automaton* $\mathcal{A}_{\varphi,\psi}$ *associated with* $\Pi(\varphi, \psi)$ is the 6-tuple

$$\mathcal{A}_{\varphi,\psi} = (Q, \Sigma_1, \Sigma_2, \delta, \iota, F),$$

where:

- $Q = Q_1 \cup Q_2 \cup \{\iota, (\varepsilon_C, \varepsilon_C)\}$, for

$$Q_1 = \{(u, \varepsilon_C) \in C^+ \times \{\varepsilon_C\} : (\exists w \in \varphi(A))(u <_s w)\},$$
$$Q_2 = \{(\varepsilon_C, v) \in \{\varepsilon_C\} \times C^+ : (\exists w \in \psi(B))(v <_s w)\};$$

- $\Sigma_1 = A$, $\Sigma_2 = B$, so that

$$\Sigma = (A \cup \{\varepsilon_A\}) \times (B \cup \{\varepsilon_B\});$$

- $\delta = \bigcup_{i=1}^{8} \Delta_i \subset Q \times \Sigma \times Q$, where

$$\Delta_1 = \{(\iota, (a, \varepsilon_B), (\varepsilon_C, \varepsilon_C)) : a \in A, \varphi(a) = \varepsilon_C\},$$
$$\Delta_2 = \{(\iota, (\varepsilon_A, b), (\varepsilon_C, \varepsilon_C)) : b \in B, \psi(b) = \varepsilon_C\},$$
$$\Delta_3 = \{(\iota, (a, b), (\psi(b)^{-1}\varphi(a), \varepsilon_C)) : a \in A, b \in B, \psi(b) \leq_p \varphi(a), \psi(b) \neq \varepsilon_C\},$$
$$\Delta_4 = \{(\iota, (a, b), (\varepsilon_C, \varphi(a)^{-1}\psi(b)) : a \in A, b \in B, \varphi(a) \leq_p \psi(b), \varphi(a) \neq \varepsilon_C\},$$
$$\Delta_5 = \{((u, \varepsilon_C), (\varepsilon_A, b), (\psi(b)^{-1}u, \varepsilon_C)) : b \in B, u \neq \varepsilon_C,\ \psi(b) \leq_p u\},$$
$$\Delta_6 = \{((u, \varepsilon_C), (\varepsilon_A, b), (\varepsilon_C, u^{-1}\psi(b))) : b \in B, u \neq \varepsilon_C,\ u \leq_p \psi(b)\},$$
$$\Delta_7 = \{((\varepsilon_C, v), (a, \varepsilon_B), (\varepsilon_C, \varphi(a)^{-1}v)) : a \in A, v \neq \varepsilon_C,\ \varphi(a) \leq_p v\},$$
$$\Delta_8 = \{((\varepsilon_C, v), (a, \varepsilon_B), (v^{-1}\varphi(a), \varepsilon_C) : a \in A, v \neq \varepsilon_C,\ v \leq_p \varphi(a)\},$$

- ι is an initial state, and $F = \{(\varepsilon_C, \varepsilon_C)\}$.

Note that there are no transitions into the initial state, and no transitions out of the final state. ■

As the above definition is written techincally, we now give two examples which illustrate how to construct the two-tape automaton associated with a fiber product.

Examples 5.2.4. (a) Let $A = \{a, b\} = B$, and let $C = \{x\}$. Define $\varphi : A^* \to C^*$ and $\psi : B^* \to C^*$ by $\varphi(a) = \varphi(b) = x$ (which is uniquely extended to an epimorphism from A^* onto C^*), and $\psi(a) = x, \psi(b) = x^2$ (which is uniquely extended to an epimorphism from B^* onto C^*).

The fiber product $\Pi(\varphi, \psi)$ can be described as

$$\Pi(\varphi, \psi) = \{(u, v) \in A^* \times B^* : |u| = |v|_a + 2|v|_b\}.$$

We now illustrate how to construct $\mathcal{A}_{\varphi,\psi}$ from $\Pi(\varphi, \psi)$. Firstly, as every state and transition is written in terms of the images of the alphabets A and B, then we note that

$$\varphi(A) = \{x\}, \psi(B) = \{x, x^2\}.$$

For the states in Q_1, the possible words $w \in \varphi(A)$ are simply $w = x$, which has no proper suffixes. Hence Q_1 is empty. For Q_2 however, as $\psi(B) = \{x, x^2\}$, then $w = x^2$ has a proper suffix x. Hence $Q_2 = \{(\varepsilon_C, x)\}$, and thus

$$Q = \{\iota, (\varepsilon_C, x), (\varepsilon_C, \varepsilon_C)\}.$$

For the transition relation δ, we consider each set Δ_i individually. Firstly, Δ_1 is empty, as $\varepsilon_C \notin \varphi(A)$. Similarly, Δ_2 is also empty, as $\varepsilon_C \notin \psi(B)$.

For Δ_3, we will obtain two edges, as $\psi(a) = x$, which is a prefix of both $\varphi(a) = x$ and $\varphi(b) = x$. As $\psi(a)^{-1}\varphi(a) = x^{-1}x = \varepsilon_C$, and

$\psi(a)^{-1}\varphi(b) = x^{-1}x = \varepsilon_C$, then

$$\Delta_3 = \{(\iota, (a, a), (\varepsilon_C, \varepsilon_C)), (\iota, (b, a), (\varepsilon_C, \varepsilon_C))\}.$$

For Δ_4, we will obtain four edges, as $\varphi(a) = x = \varphi(b)$, which are both prefixes of $\psi(a) = x$ and $\psi(b) = x^2$. These edges are

$$(\iota, (a, a), (\varepsilon_C, \varepsilon_C)),(\iota, (a, b), (\varepsilon_C, x)),$$
$$(\iota, (b, a), (\varepsilon_C, \varepsilon_C)),(\iota, (b, b), (\varepsilon_C, x)).$$

Two of these edges have already been included in δ from Δ_3, which will be accounted for when we take the union of the Δ_i.

For Δ_5, there are no states of the form (u, ε_C) for $u \neq \varepsilon_C$, hence Δ_5 is empty. The same is true for Δ_6.

For Δ_7 however, the only state of the form (ε_C, v) for $v \neq \varepsilon_C$ is (ε_C, x). As $\varphi(a) = x = \varphi(b)$ are prefixes of x, then we obtain the two edges

$$\Delta_7 = \{((\varepsilon_C, x), (a, \varepsilon_B), (\varepsilon_C, \varepsilon_C)), ((\varepsilon_C, x), (b, \varepsilon_B), (\varepsilon_C, \varepsilon_C))\}.$$

Finally, considering Δ_8, as (ε_C, x) is the only state of the form (ε_C, v) for $v \neq \varepsilon_C$, and x is a prefix of both $\varphi(a) = x$ and $\varphi(b) = x$, then we'd again obtain the two edges

$$\Delta_8 = \{((\varepsilon_C, x), (a, \varepsilon_B), (\varepsilon_C, \varepsilon_C)), ((\varepsilon_C, x), (b, \varepsilon_B), (\varepsilon_C, \varepsilon_C))\},$$

which were already accounted for in Δ_7.

Taking the union of the Δ_i as our transition relation δ, then we obtain the picture for $\mathcal{A}_{\varphi,\psi}$ given in Figure 5.1, overleaf.

(b) Let $A = \{a, b\} = B$, and let $C = \{x\}$. Define $\varphi : A^* \to C^*$ and $\psi : B^* \to C^*$ by $\varphi(a) = x$, $\varphi(b) = x^2$ (which is uniquely extended to an epimorphism from A^* onto C^*), and $\psi(a) = x^2, \psi(b) = x$ (which is uniquely extended to an epimorphism from B^* onto C^*).

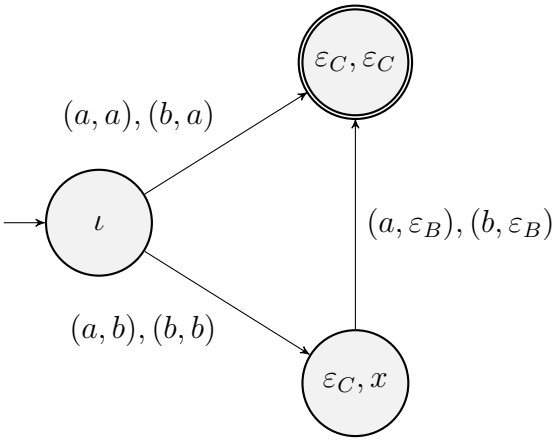


Figure 5.1: $\mathcal{A}_{\varphi,\psi}$ for $\Pi(\varphi,\psi)$ given in Examples 5.2.4 (a).

The fiber product can be given by

$$\Pi(\varphi,\psi) = \{(u,v) \in A^+ \times B^+ : |u|_a + 2|u|_b = 2|v|_a + |v|_b\}.$$

As the images of the alphabets are

$$\varphi(A) = \{x, x^2\} = \psi(B),$$

then the only proper prefixes are x, and hence we create the states (x, ε_C) according to Q_1, and (ε_C, x) according to Q_2, so that

$$Q = \{\iota, (\varepsilon_C, \varepsilon_C), (x, \varepsilon_C), (\varepsilon_C, x)\}.$$

Turning to the transition relation δ, as ε_C does not appear in either $\varphi(A)$ or $\psi(B)$, then $\Delta_1 = \Delta_2 = \emptyset$.

For Δ_3, the non-empty prefixes of words in $\varphi(A) = \{x, x^2\}$, are x or x^2, which are mapped onto by b and a via ψ, respectively. This gives us the set of three edges

$$\Delta_3 = \{(\iota, (a,b), (\varepsilon_C, \varepsilon_C)), (\iota, (b,a), (\varepsilon_C, \varepsilon_C)), (\iota, (b,b), (x, \varepsilon_C))\}.$$

For Δ_4, the non-empty prefixes of words in $\psi(B) = \{x, x^2\}$, are x or x^2,

which are mapped onto by a and b via φ, respectively. This gives us the set of three edges

$$\Delta_4 = \{(\iota, (a, b), (\varepsilon_C, \varepsilon_C)), (\iota, (b, a), (\varepsilon_C, \varepsilon_C)), (\iota, (a, a), (\varepsilon_C, x))\}.$$

For Δ_5, the only state of the form (u, ε_C) for $u \neq \varepsilon_C$ is (x, ε_C). As $\psi(b) = x$ which is a prefix of x, then we obtain the set of edges

$$\Delta_5 = \{((x, \varepsilon_C), (\varepsilon_A, b), (\varepsilon_C, \varepsilon_C))\}.$$

Similarly for Δ_6, as the only state of the form (u, ε_C) for $u \neq \varepsilon_C$ is (x, ε_C), then x is a prefix of both $\psi(b) = x$ and $\psi(a) = x^2$. This gives us the set of two edges

$$\Delta_6 = \{((x, \varepsilon_C), (\varepsilon_A, b), (\varepsilon_C, \varepsilon_C)), ((x, \varepsilon_C), (\varepsilon_A, a), (\varepsilon_C, x))\}.$$

For Δ_7, the only state of the form (ε_C, v) for $v \neq \varepsilon_C$ is (ε_C, x). As $\varphi(a) = x$ is a prefix of x, then we have the set of edges

$$\Delta_7 = \{((\varepsilon_C, x), (a, \varepsilon_B), (\varepsilon_C, \varepsilon_C))\}.$$

Finally for Δ_8, we're considering edges from the state (ε_C, x) similarly to Δ_7. As x is a prefix of both $\varphi(a) = x$ and $\varphi(b) = x^2$, then we have the set of two edges

$$\Delta_8 = \{((\varepsilon_C, x), (a, \varepsilon_B), (\varepsilon_C, \varepsilon_C)), ((\varepsilon_C, x), (b, \varepsilon_B), (x, \varepsilon_C))\}.$$

Taking δ to be the union of the Δ_i, we obtain the picture for $\mathcal{A}_{\varphi,\psi}$ given in Figure 5.2, overleaf. $\triangle$

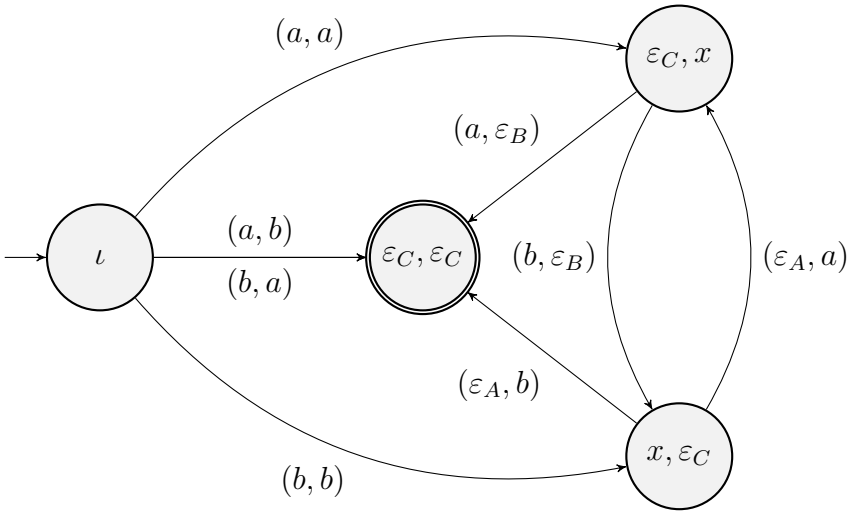


Figure 5.2: $\mathcal{A}_{\varphi,\psi}$ for $\Pi(\varphi,\psi)$ given in Examples 5.2.4 (b).

5.3 The language recognised by the two-tape automaton associated with $\Pi(\varphi,\psi)$

We now aim to show that $\mathcal{A}_{\varphi,\psi}$ as constructed in 5.2 accepts a language which corresponds to the set of indecomposable elements of $\Pi(\varphi,\psi)$. This will aid us in deciding finite generation by Lemma 5.1.2. Hence we introduce the following result.

Theorem 5.3.1. *Let $\varphi : A^* \to C^*$, $\psi : B^* \to C^*$ be two epimorphisms with A, B, C finite alphabets, and let $\mathcal{A}_{\varphi,\psi}$ be the associated automaton given in Definition* 5.2.3*. Then*

$$\pi(\mathcal{L}(\mathcal{A}_{\varphi,\psi})) = \mathcal{I},$$

where π is the product monoid mapping defined in Definition 5.2.1*, and $\mathcal{I}$ is the set of indecomposable elements of $\Pi(\varphi,\psi)$.*

Our strategy for proving Theorem 5.3.1 will be to prove each inclusion separately, by proving the following two lemmas.

Lemma 5.3.2. *For $\varphi, \psi, \mathcal{I}$ defined in Theorem 5.3.1,*

$$\pi(\mathcal{L}(\mathcal{A}_{\varphi,\psi})) \subseteq \mathcal{I}.$$

Lemma 5.3.3. *For $\varphi, \psi, \mathcal{I}$ defined in Theorem 5.3.1,*

$$\mathcal{I} \subseteq \pi(\mathcal{L}(\mathcal{A}_{\varphi,\psi})).$$

We first introduce and prove the following technical lemmas (Lemma 5.3.4 and Lemma 5.3.5) in order to prove Lemma 5.3.2 and Lemma 5.3.3 later in the section.

Lemma 5.3.4. *Let $\varphi : A^* \to C^*$, $\psi : B^* \to C^*$ be two epimorphisms with A, B, C finite alphabets, and let $\mathcal{A}_{\varphi,\psi}$ be the associated automaton given in Definition* 5.2.3.

If a path from ι to a state $(u, v) \in Q$ has label $(\alpha, \beta) \in A^* \times B^*$, then

$$\varphi(\alpha)v = \psi(\beta)u. \tag{5.1}$$

Proof. We will prove inductively on path length that (5.1) holds for any given path of length k from ι to (u, v) with label (α, β).

For the base case, the paths p of length $k = 1$ originating from ι are precisely the transitions described in Δ_1, Δ_2, Δ_3 and Δ_4. The paths $p \in \Delta_1$ ending in state $(\varepsilon_C, \varepsilon_C)$ have labels (a, ε_B) for $a \in A$, satisfying $\varphi(a) = \varepsilon_C = \psi(\varepsilon_B)$, and hence (5.1) holds. The case for Δ_2 is similar.

The paths $p \in \Delta_3$ are of the form $p = (\iota, (a, b), (\psi(b)^{-1}\varphi(a), \varepsilon_C))$ for some $a \in A$, $b \in B$. Hence the possible labels are $(\alpha, \beta) = (a, b)$, and the states are $(u, v) = (\psi(b)^{-1}\varphi(a), \varepsilon_C))$. Thus

$$\varphi(\alpha)v = \varphi(a) = \psi(b)\psi(b)^{-1}\varphi(a) = \psi(\beta)u,$$

and so (5.1) holds. Similarly, the paths $p \in \Delta_4$ are of the form $p = (\iota, (a, b), (\varepsilon_C, \varphi(a)^{-1}\psi(b))$ for some $a \in A$, $b \in B$. Hence

$$\varphi(\alpha)v = \varphi(a)\varphi(a)^{-1}\psi(b) = \psi(b) = \psi(\beta)u,$$

and so (5.1) holds. This proves the base case for paths of length $k = 1$.

Assume for the inductive hypothesis that if a path from ι to (u, v) has length k

and label (α, β), then it satisfies (5.1). Consider now a path p' of length $k+1$, from ι to (u', v') with label (α', β'). Then there is a path p from ι to a state (u, v) of length k with some label (α, β), and a transition $t \in \Delta_5 \cup \Delta_6 \cup \Delta_7 \cup \Delta_8$ from (u, v) to (u', v').

There are two possible cases for the label of the transition t, which we will consider separately:

Case 1: t has label (ε_A, b) for some $b \in B$. In this case, t belongs to either Δ_5 or Δ_6. Hence the state (u, v) is of the form (u, ε_C) for some $u \neq \varepsilon_C$.

Moreover, as the path p' has label $(\alpha', \beta') = (\alpha, \beta)(\varepsilon_A, b)$, then the path p has label $(\alpha, \beta) = (\alpha', \beta' b^{-1})$. If $\psi(b) \leq_p u$, then it follows that $t \in \Delta_5$, and hence by definition it follows that $(u', v') = (\psi(b)^{-1} u, \varepsilon_C)$. Thus

$$\begin{aligned} \psi(\beta')u' &= \psi(\beta' b^{-1})\psi(b)u' \\ &= \psi(\beta' b^{-1})u \\ &= \varphi(\alpha')v \text{ by the inductive hypothesis} \\ &= \varphi(\alpha')v'. \end{aligned}$$

Otherwise, if $u \leq_p \psi(b)$, then $t \in \Delta_6$, and it follows that $(u', v') = (\varepsilon_C, u^{-1}\psi(b))$. Hence

$$\begin{aligned} \psi(\beta')u' &= \psi(\beta' b^{-1})\psi(b) \\ &= \psi(\beta' b^{-1})uv' \\ &= \varphi(\alpha')vv' \text{ by the inductive hypothesis} \\ &= \varphi(\alpha')v'. \end{aligned}$$

In both scenarios, (5.1) is satisfied, proving the inductive step.

Case 2: t has label (a, ε_C) for some $a \in A$. This time, t belongs to either Δ_7 or Δ_8. Hence the state (u, v) is of the form (ε_C, v) for some $v \neq \varepsilon_C$.

Thus the path p' has label $(\alpha', \beta') = (\alpha, \beta)(a, \varepsilon_B)$, and hence the path p has label $(\alpha, \beta) = (\alpha' a^{-1}, \beta')$. In the case where $\varphi(a) \leq_p v$, then $t \in \Delta_7$ and

hence by definition it follows that $(u', v') = (\varepsilon_C, \varphi(a)^{-1}v)$. Thus

$$\begin{aligned}\varphi(\alpha')v' &= \varphi(\alpha' a^{-1})\varphi(a)v' \\ &= \varphi(\alpha' a^{-1})v \\ &= \psi(\beta')u \text{ by the inductive hypothesis} \\ &= \psi(\beta')u'.\end{aligned}$$

Otherwise, if $v \leq_p \varphi(a)$, then t belongs to Δ_8, and hence $(u', v') = (v^{-1}\varphi(a), \varepsilon_C)$ by the definition of Δ_8. Thus

$$\begin{aligned}\varphi(\alpha')v' &= \varphi(\alpha' a^{-1})\varphi(a) \\ &= \varphi(\alpha' a^{-1})vu' \\ &= \psi(\beta')uu' \text{ by the inductive hypothesis} \\ &= \psi(\beta')u'.\end{aligned}$$

Again, (5.1) is satisfied, proving the inductive step. Hence the result holds for paths of arbitrary length. □

Lemma 5.3.5. *For the automata $\mathcal{A}_{\varphi,\psi}$ associated with the fiber product $\Pi(\varphi, \psi)$, let $p = (q_{i-1}, \sigma_i, q_i)_{i=1}^n$ be a path of length n with $q_0 = \iota$. Define*

$$\Phi = \varphi \circ \pi_{A^*} \circ \pi, \quad \Psi = \psi \circ \pi_{B^*} \circ \pi$$

where π is the product monoid mapping defined in Definition 5.2.1*, $\pi_{A^*} : A^* \times B^* \to A^*$ is projection onto the first coordinate, and $\pi_{B^*} : A^* \times B^* \to B^*$ is projection onto the second coordinate.*

Then either

$$q_n = (\Psi(\sigma_1 \dots \sigma_n)^{-1}\Phi(\sigma_1 \dots \sigma_n), \varepsilon_C) \tag{5.2}$$

if $\Psi(\sigma_1 \dots \sigma_n) \leq_p \Phi(\sigma_1 \dots \sigma_n)$, or

$$q_n = (\varepsilon_C, \Phi(\sigma_1 \dots \sigma_n)^{-1}\Psi(\sigma_1 \dots \sigma_n)) \tag{5.3}$$

if $\Phi(\sigma_1 \dots \sigma_n) \leq_p \Psi(\sigma_1 \dots \sigma_n)$.

Proof. We will proceed by induction on n. For the base case where $n = 1$, the path p is a transition $(\iota, \sigma_1, q_1) \in \Delta_1 \cup \Delta_2 \cup \Delta_3 \cup \Delta_4$.

If $p \in \Delta_1$, then $\sigma_1 = (a, \varepsilon_B)$ for some $a \in A$ with $\varphi(a) = \varepsilon_C$. Hence

$$\Phi(\sigma_1) = (\varphi \circ \pi_{A^*} \circ \pi)(a, \varepsilon_C) = \varphi(a) = \varepsilon_C, \quad \Psi(\sigma_1) = (\psi \circ \pi_{B^*} \circ \pi)(a, \varepsilon_C) = \varepsilon_C,$$

and $q_1 = (\varepsilon_C, \varepsilon_C) = (\Psi(\sigma_1)^{-1}\Phi(\sigma_1), \varepsilon_C)$ as expected. The case for $p \in \Delta_2$ is almost identical.

For $p \in \Delta_3$ and $p \in \Delta_4$, (5.2) and (5.3) follow by definition of Δ_3 and Δ_4, respectively. This proves the base case.

Assume for the inductive hypothesis that for all paths $p = (q_{i-1}, \sigma_i, q_i)_{i=1}^{k}$ of length k, either

$$q_k = (\Psi(\sigma_1 \dots \sigma_k)^{-1}\Phi(\sigma_1 \dots \sigma_k), \varepsilon_C), \tag{5.4}$$

if $\Psi(\sigma_1 \dots \sigma_k) \leq_p \Phi(\sigma_1 \dots \sigma_k)$, or

$$q_k = (\varepsilon_C, \Phi(\sigma_1 \dots \sigma_k)^{-1}\Psi(\sigma_1 \dots \sigma_k)), \tag{5.5}$$

if $\Phi(\sigma_1 \dots \sigma_k) \leq_p \Psi(\sigma_1 \dots \sigma_k)$.

Consider the state q'_{k+1} for a $k+1$ length path $p' = (q'_{i-1}, \sigma'_i, q'_i)_{i=1}^{k+1}$ in the case where $q'_k \neq (\varepsilon_C, \varepsilon_C)$. By the inductive hypothesis, state q'_k satisfies either (5.4) or (5.5).

In the case of (5.4), then as we assume $\Psi(\sigma'_1 \dots \sigma'_k)^{-1}\Phi(\sigma'_1 \dots \sigma'_k) \neq \varepsilon_C$, the transition $(q'_k, \sigma'_{k+1}, q'_{k+1})$ is either an element of Δ_5 or Δ_6. Hence $\sigma'_{k+1} = (\varepsilon_A, b)$ for some $b \in B$.

If $\psi(b) \leq_p \Psi(\sigma'_1 \dots \sigma'_k)^{-1}\Phi(\sigma'_1 \dots \sigma'_k)$, then by the definition of Δ_5, we have

$$\begin{aligned} q'_{k+1} &= (\psi(b)^{-1}\Psi(\sigma'_1 \dots \sigma'_k)^{-1}\Phi(\sigma'_1 \dots \sigma'_k), \varepsilon_C) \\ &= ([\Psi(\sigma'_1 \dots \sigma'_k)\psi(b)]^{-1}\Phi(\sigma'_1 \dots \sigma'_k), \varepsilon_C) \\ &= ([\Psi(\sigma'_1 \dots \sigma'_k)\Psi(\sigma'_{k+1})]^{-1}\Phi(\sigma'_1 \dots \sigma'_k), \varepsilon_C) \\ &= (\Psi(\sigma'_1 \dots \sigma'_{k+1})^{-1}\Phi(\sigma'_1 \dots \sigma'_{k+1}), \varepsilon_C), \end{aligned}$$

and so 5.2 holds. Otherwise if $\Psi(\sigma'_1 \dots \sigma'_k)^{-1}\Phi(\sigma'_1 \dots \sigma'_k) \leq_p \psi(b)$, then by the definition of Δ_6,

$$\begin{aligned} q'_{k+1} &= (\varepsilon_C, [\Psi(\sigma'_1 \dots \sigma'_k)^{-1}\Phi(\sigma'_1 \dots \sigma'_k)]^{-1}\psi(b)) \\ &= (\varepsilon_C, \Phi(\sigma'_1 \dots \sigma'_k)^{-1}\Psi(\sigma'_1 \dots \sigma'_k)\psi(b)) \\ &= (\varepsilon_C, \Phi(\sigma'_1 \dots \sigma'_k)^{-1}\Psi(\sigma'_1 \dots \sigma'_k)\Psi(\sigma'_{k+1})) \\ &= (\varepsilon_C, \Phi(\sigma'_1 \dots \sigma'_{k+1})^{-1}\Psi(\sigma'_1 \dots \sigma'_{k+1})), \end{aligned}$$

and so 5.3 holds.

The case for (5.5) can be proven symmetrically. Hence the inductive step has been proved, and the result follows for arbitrary n. □

We are now in a position to prove Theorem 5.3.1 by proving Lemma 5.3.2 and Lemma 5.3.3 which we restate and prove below.

Lemma 5.3.2. *For $\varphi, \psi, \mathcal{I}$ defined in Theorem 5.3.1,*

$$\pi(\mathcal{L}(\mathcal{A}_{\varphi,\psi})) \subseteq \mathcal{I}.$$

Proof. Let $\sigma_1\sigma_2 \dots \sigma_k \in \mathcal{L}(\mathcal{A}_{\varphi,\psi})$ be a word of length k, and consider the image
$\pi(\sigma_1\sigma_2 \dots \sigma_k) = (\alpha, \beta) \in A^* \times B^*$.

As $\sigma_1\sigma_2 \dots \sigma_k \in L$, there exists a length k path $p = (q_{i-1}, \sigma_i, q_i)_{i=1}^k$ in $\mathcal{A}_{\varphi,\psi}$, from $q_0 = \iota$ to $(\varepsilon_C, \varepsilon_C)$ with label (α, β). By Lemma 5.3.4, it follows that $\varphi(\alpha) = \psi(\beta)$, and hence $(\alpha, \beta) \in \Pi(\varphi, \psi)$.

Suppose for a contradiction that (α, β) were decomposable in $\Pi(\varphi, \psi)$. Then there would exist $(\alpha', \beta'), (\alpha'', \beta'') \in \Pi(\varphi, \psi) \setminus \{(\varepsilon_A, \varepsilon_B)\}$ such that

$$(\alpha, \beta) = (\alpha', \beta')(\alpha'', \beta'') \tag{5.6}$$

is a non-trivial decomposition of (α, β).

By the definition of the transition relation δ in $\mathcal{A}_{\varphi,\psi}$, either $\sigma_1 = (a, \varepsilon_B)$, $\sigma_1 = (\varepsilon_A, b)$, or $\sigma_1 = (a, b)$ for some $a \in A$, $b \in B$. The first two situations

lead to a contradiction, as either $p = (\iota, (a, \varepsilon_B), (\varepsilon_C, \varepsilon_C))$ or $p = (\iota, (\varepsilon_A, b), (\varepsilon_C, \varepsilon_C))$, and hence either $(\alpha, \beta) = (a, \varepsilon_B)$ or $(\alpha, \beta) = (\varepsilon_A, b)$ which are indecomposable.

Thus $\sigma_1 = (a, b)$ for some $a \in A$, $b \in B$. In particular, it must be that $\sigma_1 = (|\alpha|_1, |\beta|_1)$ (and moreover, $(\alpha, \beta) \in A^+ \times B^+$).

Suppose $\alpha' = \varepsilon_A$. Then as $(\alpha', \beta') \in \Pi(\varphi, \psi) \setminus \{(\varepsilon_A, \varepsilon_B)\}$, it must be that $\beta' \in B^+$ and $\psi(\beta') = \varepsilon_C$. Writing $\beta' = b_1 b_2 \dots b_i$ for some $b_1, b_2, \dots, b_i \in B$, it would follow that

$$\psi(\beta') = \psi(b_1 b_2 \dots b_i) = \psi(b_1)\psi(b_2) \dots \psi(b_i) = \varepsilon_C,$$

and hence in particular $\psi(b_1) = \varepsilon_C$. This is a contradiction, as $\psi(|\beta|_1) = \psi(|\beta'|_1) = \psi(b_1) = \varepsilon_C$ and $\sigma_1 = (|\alpha|_1, |\beta|_1)$, but there are no transitions in δ of the form $(\iota, (|\alpha|_1, |\beta|_1), q_1)$ where $\psi(|\beta|_1) = \varepsilon_C$. Hence $\alpha' \neq \varepsilon_A$

A similar proof shows that $\beta' \neq \varepsilon_B$, and hence it must be that $(\alpha', \beta') \in A^+ \times B^+$. Letting $(\alpha, \beta) = (a_1 a_2 \dots a_{|\alpha|}, b_1 b_2 \dots b_{|\beta|})$, then

$$(\alpha', \beta') = (a_1 a_2 \dots a_m, b_1 b_2 \dots b_n)$$

for some $1 \leq m \leq |\alpha|$, $1 \leq n \leq |\beta|$ with $m+n < |\alpha|+|\beta|$. As $a_1 a_2 \dots a_m$ and $\pi_{A^*}(\pi(\sigma_1 \sigma_2 \dots \sigma_i))$ (for $1 \leq i \leq k$) are prefixes of α, then there is a minimal $M \in \mathbb{N}$ such that $a_1 a_2 \dots a_m = \pi_{A^*}(\pi(\sigma_1 \sigma_2 \dots \sigma_M))$. Similarly, we can choose an $N \in \mathbb{N}$ minimally such that $b_1 b_2 \dots b_n = \pi_{B^*}(\pi(\sigma_1 \sigma_2 \dots \sigma_N))$.

It cannot be that the length k of path p is equal to one, for then $(\alpha, \beta) = \pi(\sigma_1) = (a, b)$ for some $a \in A, b \in B$, with either $\varphi(a) \neq \varepsilon_C$ or $\psi(b) \neq \varepsilon_C$, and hence by definition of δ it would follow that (α, β) is indecomposable.

Thus $k > 1$, and further by definition of δ, each σ_i for $i > 1$ is either equal to (a, ε_B) or (ε_A, b) for some $a \in A$, $b \in B$, and so it cannot be that $M = N$.

Letting $\mu = \min\{M, N\}$, it then follows that

$$\pi(\sigma_1\sigma_2\ldots\sigma_\mu) = \begin{cases} (a_1a_2\ldots a_m, b_1b_2\ldots b_{n'}) & \text{for some } n' < n \text{ if } M < N, \\ (a_1a_2\ldots a_{m'}, b_1b_2\ldots b_n) & \text{for some } m' < m \text{ if } N < M. \end{cases}$$

In the first case where $\mu = M$, we claim that

$$\pi(\sigma_1\sigma_2\ldots\sigma_{M+t}) = (a_1a_2\ldots a_m, b_1b_2\ldots b_{n'+t}), \tag{5.7}$$

for all $0 \leq t \leq n - n'$. We will show this by proving the case for $t = 0$, and then showing that the case for $t = l$ implies the case for $t = l + 1$, when $l < n - n'$.

For the first case where $t = 0$, we have already seen that

$$\pi(\sigma_1\sigma_2\ldots\sigma_M) = (a_1a_2\ldots a_m, b_1b_2\ldots b_{n'}).$$

If the case for $t = l$ is true, then

$$\pi(\sigma_1\sigma_2\ldots\sigma_{M+l}) = (a_1a_2\ldots a_m, b_1b_2\ldots b_{n'+l}),$$

For the next case $t = l+1$, we will determine σ_{M+l+1} by considering the state q_{M+l}. As $\psi(b_1b_2\ldots b_{n'+l}) \leq_p \psi(b_1b_2\ldots b_n)$ for $l < n - n'$, and $\psi(b_1b_2\ldots b_n) = \varphi(a_1a_2\ldots a_m)$, then

$$\psi(b_1b_2\ldots b_l) \leq_p \varphi(a_1a_2\ldots a_m). \tag{5.8}$$

Recalling the definitions of Φ and Ψ from Lemma 5.3.5, as

$$\Phi(\sigma_1\sigma_2\ldots\sigma_{M+l}) = (\varphi \circ \pi_{A^*} \circ \pi)(\sigma_1\sigma_2\ldots\sigma_{M+l}) = \varphi(a_1a_2\ldots a_m)$$

and

$$\Psi(\sigma_1\sigma_2\ldots\sigma_{M+l}) = (\psi \circ \pi_{B^*} \circ \pi)(\sigma_1\sigma_2\ldots\sigma_{M+l}) = \psi(b_1b_2\ldots b_{n'+l}),$$

then by (5.8) and Lemma 5.3.5, the state q_{M+l} is either in Q_1 or $q_M = (\varepsilon_C, \varepsilon_C)$. The latter case leads to a contradiction, as then

$(\alpha, \beta) = (a_1 a_2 \dots a_m, b_1 b_2 \dots b_{n'+l})$ with $n' + l < n \leq |\beta|$. Hence it must be that $q_{M+l} \in Q_1$, and hence $\sigma_{M+l+1} = (\varepsilon_A, b)$ for some $b \in B$.

As

$$\pi_{B^*}(\pi(\sigma_1 \sigma_2 \dots \sigma_{M+l} \sigma_{M+l+1})) = b_1 b_2 \dots b_{n'+l} b,$$

and $\pi_{B^*}(\pi(\sigma_1 \sigma_2 \dots \sigma_M \sigma_{M+l+1}))$ is a prefix of β', then it must be that $b = b_{n'+l+1}$. Thus as $\pi_{A^*}(\pi(\sigma_1 \sigma_2 \dots \sigma_M \sigma_{M+l+1})) = (a_1 a_2 \dots a_m)$, we have shown that

$$\pi(\sigma_1 \sigma_2 \dots \sigma_{M+l+1}) = (a_1 a_2 \dots a_m, b_1 b_2 \dots b_{n'+l+1}),$$

and have hence proved our claim in (5.7), as $t = 0$ implies $t = 1$, and $t = 1$ implies $t = 2$ and so on, until $t = n - n'$. Thus in the case that $\mu = M$, in particular we have

$$\pi(\sigma_1 \sigma_2 \dots \sigma_{M+n-n'}) = (a_1 a_2 \dots a_m, b_1 b_2 \dots b_n).$$

Hence by Lemma 5.3.5, it follows that $q_{M+n-n'} = (\varepsilon_C, \varepsilon_C)$, as $(a_1 a_2 \dots a_m, b_1 b_2 \dots b_n) \in \Pi(\varphi, \psi)$. As there are no transitions leading out from the state $(\varepsilon_C, \varepsilon_C)$ by the definition of δ, it must then be that $q_k = q_{M+n-n'}$, and thus $(\alpha, \beta) = (\alpha', \beta')$.

This is a contradiction of the non-triviality of the decomposition in (5.6), and hence (α, β) is indecomposable in the case where $\mu = M$. The case for $\mu = N$ is similar to $\mu = M$, when showing that

$$\pi(\sigma_1 \sigma_2 \dots \sigma_{N+t}) = (a_1 a_2 \dots a_{m'+t}, b_1 b_2 \dots b_n)$$

for $0 \leq t \leq m' - m$, and deducing that $q_{N+m-m'} = (\varepsilon_C, \varepsilon_C)$, again showing by Lemma 5.3.5 that $(\alpha, \beta) = (\alpha', \beta')$.

Thus we have shown that $(\alpha, \beta) \in I$ for any $(\alpha, \beta) \in \pi(\mathcal{L}(\mathcal{A}_{\varphi,\psi}))$, and hence the result follows. □

Lemma 5.3.3. *For $\varphi, \psi, \mathcal{I}$ defined in Theorem 5.3.1,*

$$\mathcal{I} \subseteq \pi(\mathcal{L}(\mathcal{A}_{\varphi,\psi})).$$

Proof. Let $(\alpha, \beta) \in \Pi(\varphi, \psi)$ be an indecomposable element. We will construct a path $p = (q_{i-1}, \sigma_i, q_i)_{i=1}^{k}$ with $q_0 = \iota$ and $q_k = (\varepsilon_C, \varepsilon_C)$ so that $\sigma_1\sigma_2 \dots \sigma_k \in \mathcal{L}(\mathcal{A}_{\varphi,\psi})$, and $\pi(\sigma_1\sigma_2 \dots \sigma_k) = (\alpha, \beta)$.

In the particular case where $\alpha = \varepsilon_A$, it must be that $\beta \in B^+$. Writing $\beta = b_1b_2 \dots b_n$ for some $b_1, b_2, \dots b_n \in B$, then as

$$\varepsilon_C = \varphi(\alpha) = \psi(\beta) = \psi(b_1b_2 \dots b_n) = \psi(b_1)\psi(b_2) \dots \psi(b_n),$$

it follows that $\psi(b_i) = \varepsilon_C$ for each $1 \leq i \leq n$. Hence if (α, β) is indecomposable with $\alpha = \varepsilon_A$, it follows that $(\alpha, \beta) = (\varepsilon_A, b)$ for some $b \in B$. Now the transition

$$p = (\iota, (\varepsilon_A, b), (\varepsilon_C, \varepsilon_C)) \in \Delta_2$$

is a path accepting the word $\sigma_1 = (\varepsilon_A, b) \in \Sigma^*$, with $\pi(\sigma_1) = (\alpha, \beta)$, and hence $(\alpha, \beta) \in \pi(\mathcal{L}(\mathcal{A}_{\varphi,\psi}))$ in this particular case.

Similarly, if $\beta = \varepsilon_B$, then $(\alpha, \beta) = (a, \varepsilon_B)$ for some $a \in A$, and the transition $p = (\iota, (a, \varepsilon_B), (\varepsilon_C, \varepsilon_C)) \in \Delta_1$ is a path accepting the word $\sigma_1 = (a, \varepsilon_B) \in \Sigma^*$, with $\pi(\sigma_1) = (\alpha, \beta)$. Hence again, $(\alpha, \beta) \in \pi(\mathcal{L}(\mathcal{A}_{\varphi,\psi}))$.

We will thus consider the final case where $(\alpha, \beta) \in A^+ \times B^+$. Hence suppose

$$(\alpha, \beta) = (a_1a_2 \dots a_{|\alpha|}, b_1b_2 \dots b_{|\beta|})$$

for some $a_1, a_2, \dots, a_{|\alpha|} \in A$, $b_1, b_2, \dots, b_{|\beta|} \in B$. Define the sequence of triples $p = (q_{i-1}, \sigma_i, q_i)_{i=1}^{|\alpha|+|\beta|-1} \in Q \times \Sigma \times Q$ (where Q, Σ are as in Definition 5.2.3) by $q_0 = \iota$, $\sigma_1 = (a_1, b_1)$, and (recalling the definitions of Φ and Ψ from Lemma 5.3.5)

$$q_i = \begin{cases} (\Psi(\sigma_1 \dots \sigma_i)^{-1}\Phi(\sigma_1 \dots \sigma_i), \varepsilon_C) & \text{if } \Psi(\sigma_1 \dots \sigma_i) \leq_p \Phi(\sigma_1 \dots \sigma_i) \\ (\varepsilon_C, \Phi(\sigma_1 \dots \sigma_i)^{-1}\Psi(\sigma_1 \dots \sigma_i)) & \text{if } \Phi(\sigma_1 \dots \sigma_i) \leq_p \Psi(\sigma_1 \dots \sigma_i) \end{cases}$$

for $1 \leq i \leq |\alpha| + |\beta| - 1$, and finally

$$\sigma_i = \begin{cases} (a_{j_{i-1}}, \varepsilon_B) & \text{if } q_{i-1} \in Q_2 \\ (\varepsilon_A, b_{k_{i-1}}) & \text{if } q_{i-1} \in Q_1 \end{cases} \tag{5.9}$$

(where $j_{i-1} = |\pi_{A^*}(\pi(\sigma_1 \dots \sigma_{i-1}))| + 1$, and $k_{i-1} = |\pi_{B^*}(\pi(\sigma_1 \dots \sigma_{i-1}))| + 1$ for $2 \leq i \leq |\alpha| + |\beta| - 1$).

Firstly, each q_i is a well defined state in $\mathcal{A}_{\varphi,\psi}$, as for each $1 \leq i \leq |\alpha| + |\beta| - 1$, $\Phi(\sigma_1\sigma_2 \dots \sigma_i)$ is a prefix of $\varphi(\alpha)$, $\Phi(\sigma_1\sigma_2 \dots \sigma_i)$ is a prefix of $\psi(\beta)$, and $\varphi(\alpha) = \psi(\beta)$. Hence either $\Phi(\sigma_1\sigma_2 \dots \sigma_i)$ is a prefix of $\Psi(\sigma_1\sigma_2 \dots \sigma_i)$, or vice versa for each i.

Moreover, $q_i \neq (\varepsilon_C, \varepsilon_C)$ for $i < |\alpha| + |\beta| - 1$ by indecomposability of (α, β). Hence as each $q_i \in Q_1 \cup Q_2$ for $1 \leq i \leq |\alpha| + |\beta| - 1$, then each σ_i is a well defined element of Σ in $\mathcal{A}_{\varphi,\psi}$.

We make the following claims about the sequence of triples $p = (q_{i-1}, \sigma_i, q_i)_{i=1}^{|\alpha|+|\beta|-1}$ just defined:

Claim 1: $(q_{i-1}, \sigma_i, q_i) \in \delta$ *for* $1 \leq i \leq |\alpha| + |\beta| - 1$, *and hence* p *is a path in* $\mathcal{A}_{\varphi,\psi}$.

Claim 2: $\pi(\sigma_1\sigma_2 \dots \sigma_{|\alpha|+|\beta|-1}) = (\alpha, \beta)$.

From Claim 2, it follows that $q_{|\alpha|+|\beta|-1} = (\varepsilon_C, \varepsilon_C)$ by Lemma 5.3.5 as

$$\Phi(\sigma_1\sigma_2 \dots \sigma_{|\alpha|+|\beta|-1}) = \varphi(\alpha) = \psi(\beta) = \Psi(\sigma_1\sigma_2 \dots \sigma_{|\alpha|+|\beta|-1}),$$

and thus $\sigma_1\sigma_2 \dots \sigma_{|\alpha|+|\beta|-1} \in \mathcal{L}(\mathcal{A}_{\varphi,\psi})$. Hence $(\alpha, \beta) \in \pi(\mathcal{L}(\mathcal{A}_{\varphi,\psi}))$, completing the proof of the result.

It thus remains to prove Claim 1 and Claim 2.

Proof of Claim 1: In the case of $i = 1$, $\sigma_1 = (a_1, b_1)$ by definition. As $\Phi(\sigma_1) = \varphi(a_1)$, and $\Psi(\sigma_1) = \psi(b_1)$, then as $\varphi(\alpha) = \psi(\beta)$, either $\psi(b_1) \leq_p \varphi(a_1)$ or $\varphi(a_1) \leq_p \psi(b_1)$. Hence either

$$q_1 = (\Psi(\sigma_1)^{-1}\Phi(\sigma_1), \varepsilon_C) = (\psi(b_1)^{-1}\varphi(a_1), \varepsilon_C)$$

if $\psi(b_1) \leq_p \varphi(a_1)$, thus $(q_0, \sigma_1, q_1) = (\iota, (a_1, b_1), (\psi(b_1)^{-1}\varphi(a_1), \varepsilon_C)) \in \Delta_3$ as required, or

$$q_1 = (\varepsilon_C, \Phi(\sigma_1)^{-1}\Psi(\sigma_1)^{-1}) = (\varepsilon_C, \varphi(a_1)^{-1}\psi(b_1))$$

if $\varphi(a_1) \leq_p \psi(b_1)$, thus $(q_0, \sigma_1, q_1) = (\iota, (a_1, b_1), (\varepsilon_C, \varphi(a_1)^{-1}\psi(b_1)) \in \Delta_4$ as required.

For $i > 1$, the triple (q_{i-1}, σ_i, q_i) is such that $q_{i-1} \in Q_1 \cup Q_2$. We will consider the cases where $q_{i-1} \in Q_1$ and $q_{i-1} \in Q_2$ separately, showing in both cases that $(q_{i-1}, \sigma_i, q_i) \in \delta$ for $1 < i \leq |\alpha| + |\beta| - 1$.

If $q_{i-1} \in Q_1$, then $q_{i-1} = (\Psi(\sigma_1 \dots \sigma_{i-1})^{-1}\Phi(\sigma_1 \dots \sigma_{i-1}), \varepsilon_C)$ and $\sigma_i = (\varepsilon_A, b_{k_{i-1}})$ by construction, where $k_{i-1} = |\pi_{B^*}(\pi(\sigma_1\sigma_2 \dots \sigma_{i-1}))| + 1$. If $\Psi(\sigma_1 \dots \sigma_i) \leq_p \Phi(\sigma_1 \dots \sigma_i)$, then as $\Phi(\sigma_1 \dots \sigma_i) = \Phi(\sigma_1 \dots \sigma_{i-1})$, and $\Psi(\sigma_1 \dots \sigma_i) = \Psi(\sigma_1 \dots \sigma_{i-1})\psi(b_{k_{i-1}})$, it follows that $\psi(b_{k_{i-1}})$ is a prefix of $\Psi(\sigma_1 \dots \sigma_{i-1})^{-1}\Phi(\sigma_1 \dots \sigma_{i-1})$. Moreover, by construction,

$$\begin{aligned} q_i &= (\Psi(\sigma_1 \dots \sigma_i)^{-1}\Phi(\sigma_1 \dots \sigma_i), \varepsilon_C) \\ &= ([\Psi(\sigma_1 \dots \sigma_{i-1})\Psi(\sigma_i)]^{-1}\Phi(\sigma_1 \dots \sigma_{i-1}), \varepsilon_C) \\ &= (\Psi(\sigma_i)^{-1}\Psi(\sigma_1 \dots \sigma_{i-1})^{-1}\Phi(\sigma_1 \dots \sigma_{i-1}), \varepsilon_C) \\ &= (\psi(b_{k_i})^{-1}\Psi(\sigma_1 \dots \sigma_{i-1})^{-1}\Phi(\sigma_1 \dots \sigma_{i-1}), \varepsilon_C). \end{aligned}$$

Hence $(q_{i-1}, \sigma_i, q_i) \in \Delta_5 \subseteq \delta$ by definition. Similarly if $\Phi(\sigma_1 \dots \sigma_i) \leq_p \Psi(\sigma_1 \dots \sigma_i)$, then as $\Phi(\sigma_1 \dots \sigma_i) = \Phi(\sigma_1 \dots \sigma_{i-1})$ and $\Psi(\sigma_1 \dots \sigma_i) = \Psi(\sigma_1 \dots \sigma_{i-1})\psi(b_{k_{i-1}})$, it follows that $\Psi(\sigma_1 \dots \sigma_{i-1})^{-1}\Phi(\sigma_1 \dots \sigma_{i-1})$ is a prefix of $\psi(b_{k_{i-1}})$. Moreover, by construction,

$$\begin{aligned} q_i &= (\varepsilon_C, \Phi(\sigma_1 \dots \sigma_i)^{-1}\Psi(\sigma_1 \dots \sigma_i)) \\ &= (\varepsilon_C, \Phi(\sigma_1 \dots \sigma_{i-1})^{-1}\Psi(\sigma_1 \dots \sigma_{i-1})\Psi(\sigma_i)) \\ &= (\varepsilon_C, [\Psi(\sigma_1 \dots \sigma_{i-1})^{-1}\Phi(\sigma_1 \dots \sigma_{i-1})]^{-1}\Psi(\sigma_i)) \\ &= (\varepsilon_C, [\Psi(\sigma_1 \dots \sigma_{i-1})^{-1}\Phi(\sigma_1 \dots \sigma_{i-1})]^{-1}\psi(b_{k_i})), \end{aligned}$$

thus $(q_{i-1}, \sigma_i, q_i) \in \Delta_6 \subseteq \delta$ by definition.

Otherwise, if $q_{i-1} \in Q_2$, then $q_{i-1} = (\varepsilon_C, \Phi(\sigma_1 \dots \sigma_{i-1})^{-1}\Psi(\sigma_1 \dots \sigma_{i-1}))$ and $\sigma_i = (a_{j_{i-1}}, \varepsilon_B)$ by construction, where $j_{i-1} = |\pi_{A^*}(\pi(\sigma_1\sigma_2 \dots \sigma_{i-1}))| + 1$. A similar proof to the case where $q_{i-1} \in Q_1$ gives that $(q_{i-1}, \sigma_i, q_i) \in \Delta_7 \subseteq \delta$ if $\Phi(\sigma_1 \dots \sigma_i) \leq_p \Psi(\sigma_1 \dots \sigma_i)$, and that $(q_{i-1}, \sigma_i, q_i) \in \Delta_8 \subseteq \delta$ if

$\Psi(\sigma_1 \dots \sigma_i) \leq_p \Phi(\sigma_1 \dots \sigma_i)$.

Thus having considered all cases, we have shown that $(q_{i-1}, \sigma_i, q_i) \in \delta$ for $1 \leq i \leq |\alpha| + |\beta| - 1$, concluding the proof of Claim 1.

Proof of Claim 2: We will show that $\pi(\sigma_1\sigma_2 \dots \sigma_{|\alpha|+|\beta|-1}) = (\alpha, \beta)$. Firstly, we will consider the possible values for each j_i, k_i as defined in the construction of the σ_i for $1 \leq i \leq |\alpha| + |\beta| - 1$.

In the case of j_1, as $\sigma_1 = (a_1, b_1)$ by definition, then

$$j_1 = |\pi_{A^*}(\pi(\sigma_1))| + 1 = |a_1| + 1 = 2.$$

More generally for $i > 1$, if $\sigma_i = (a_{j_{i-1}}, \varepsilon_B)$, then

$$j_i = |\pi_{A^*}(\sigma_1 \dots \sigma_i)| + 1 = (|\pi_{A^*}(\sigma_1 \dots \sigma_{i-1})| + 1) + 1 = j_{i-1} + 1,$$

whereas if $\sigma_i = (\varepsilon_A, b_{k_{i-1}})$, then

$$j_i = |\pi_{A^*}(\sigma_1 \dots \sigma_i)| + 1 = |\pi_{A^*}(\sigma_1 \dots \sigma_{i-1})| + 1 = j_{i-1}.$$

It follows that the set

$$S_A = \{j_i : 2 \leq i \leq |\alpha| + |\beta| - 1, \sigma_i \in A \times \{\varepsilon_B\}\}$$

is equal to $\{2, 3, 4 \dots, m\}$ for some $m \leq |\alpha|$, and moreover satisfies $j_i < j_{i'} \Leftrightarrow i < i'$. Now as

$$\pi_{A^*}(\pi(\sigma_1\sigma_2 \dots \sigma_{|\alpha|+|\beta|-1})) = a_1 a_{j_{i_1}} a_{j_{i_2}} \dots a_{j_{i_\mu}}$$

for $j_{i_1}, j_{i_2}, \dots j_{i_\mu} \in S_A$ with $\mu = |S_A|$ and $i_1 < i_2 < \dots < i_\mu$, it must be that

$$\pi_{A^*}(\pi(\sigma_1\sigma_2 \dots \sigma_{|\alpha|+|\beta|-1})) = a_1 a_2 a_3 \dots a_m. \tag{5.10}$$

A very similar proof shows that when considering the values of k_i, the set

$$S_B = \{k_i : 1 \leq i \leq |\alpha| + |\beta| - 1, \sigma_i \in \{\varepsilon_A\} \times B\}$$

is equal to $\{2, 3, 4 \ldots, n\}$ for some $n \leq |\beta|$, satisfies $k_i < k_{i'} \Leftrightarrow i < i'$, and hence we can similarly deduce that

$$\pi_{B^*}(\pi(\sigma_1\sigma_2 \ldots \sigma_{|\alpha|+|\beta|-1})) = b_1b_2b_3 \ldots b_n. \tag{5.11}$$

Together, (5.10) and (5.11) imply that

$$\pi(\sigma_1\sigma_2 \ldots \sigma_{|\alpha|+|\beta|-1}) = (a_1a_2a_3 \ldots a_m, b_1b_2b_3 \ldots b_n),$$

and hence to prove the claim it remains to show that $m = |\alpha|$, and $n = |\beta|$.

As $m \leq |\alpha|$, suppose for a contradiction that $m < |\alpha|$. As $m = |S_A| + 1$, it follows that $|S_A| < |\alpha| - 1$. As

$$|S_A| + |S_B| = |\alpha| - 1 + |\beta| - 1 = |\alpha| + |\beta| - 2,$$

it now follows that

$$|\alpha| + |\beta| - 2 < |\alpha| - 1 + |S_B| \Rightarrow |\beta| - 1 < |S_B|.$$

But as $|S_B| = n - 1$, it follows that $|\beta| - 1 < n - 1$, and hence $|\beta| < n$ which is a contradiction, as we saw that $n \leq |\beta|$. Hence it must be that $m = |\alpha|$. A similar proof shows that $n = |\beta|$, and hence we have proved Claim 2, and thus the result. □

By proving Lemma 5.3.2 and Lemma 5.3.3, we have now proved Theorem 5.3.1, which we restate below.

Theorem 5.3.1. *Let $\varphi : A^* \to C^*$, $\psi : B^* \to C^*$ be two epimorphisms with A, B, C finite alphabets, and let $\mathcal{A}_{\varphi,\psi}$ be the associated automaton given in Definition* 5.2.3. *Then*

$$\pi(\mathcal{L}(\mathcal{A}_{\varphi,\psi})) = \mathcal{I},$$

where π is the product monoid mapping defined in Definition 5.2.1*, and $\mathcal{I}$ is the set of indecomposable elements of $\Pi(\varphi, \psi)$.*

5.4 Decidability of finite generation for fiber products of free semigroups and monoids over free fibers

Having constructed automata recognising a language corresponding to the indecomposable elements of the fiber product $\Pi(\varphi, \psi)$, in this section we will use the automata to determine finite generation of $\Pi(\varphi, \psi)$. We begin by presenting the following result as a corollary to Theorem 5.3.1, which gives a necessary and sufficient criterion for the fiber product of two free monoids over a free monoid fiber to be finitely generated.

Theorem 5.4.1. *Let $\varphi : A^* \to C^*$, $\psi : B^* \to C^*$ be two epimorphisms with A, B, C finite alphabets, and let $\mathcal{A}_{\varphi,\psi}$ be the associated automaton given in Definition* 5.2.3. *Then the following are equivalent:*

(i) $\Pi(\varphi, \psi)$ *is finitely generated;*

(ii) $\mathcal{A}_{\varphi,\psi}$ *is acyclic.*

Proof. For (i) $\Rightarrow$ (ii), we will prove the contrapositive: if $\mathcal{A}_{\varphi,\psi}$ has a cycle, then $\Pi(\varphi, \psi)$ is not finitely generated.

Hence let $(q_{i-1}, \sigma_i, q_i)_{i=1}^k$ be a cycle in $\mathcal{A}_{\varphi,\psi}$, so that $q_0 = q_k$. As there are no transitions of the form $((\varepsilon_C, \varepsilon_C), \sigma_i, q_i)$ or $(q_{i-1}, \sigma_i, \iota)$ in δ by definition, then it must be that $q_0 \in Q_1 \cup Q_2$. That is, either $q_0 = (u, \varepsilon_C)$ for some $u \in C^+$ with $u <_s w$ for some $w \in \varphi(A)$, or $q_0 = (\varepsilon_C, v)$ for some $v \in C^+$ with $v <_s w'$ for some $w' \in \psi(B)$.

We will consider the case for $q_0 = (u, \varepsilon_C)$, as the case for $q_0 = (\varepsilon_C, v)$ can be given by a similar symmetric argument. As u is a suffix of $w = \varphi(a)$ for some $a \in A$, and ψ is surjective, then there exist $b_1, \ldots, b_j, b'_1, \ldots, b'_l \in B$ such that $\psi(b_1 \ldots b_j)u = w$ and $\psi(b'_1 \ldots b'_l) = u$. In particular, each $b_i, b'_i \in B$ can be chosen so that $\psi(b_i) \neq \varepsilon_C$ and $\psi(b'_i) \neq \varepsilon_C$.

We construct the sequences of transitions $(p_{i-1}, \tau_i, p_i)_{i=1}^j$ and $(r_{i-1}, \rho_i, r_i)_{i=1}^l$ where

1. $(p_0, \tau_1, p_1) = (\iota, (a, b_1), (\psi(b_1)^{-1}w, \varepsilon_C))$,

2. $p_i = (\psi(b_1...b_i)^{-1}w, \varepsilon_C)$, $\tau_i = (\varepsilon_A, b_i)$ for $1 < i \leq j$,

3. $r_0 = q_0$, $\rho_i = (\varepsilon_A, b'_i)$, $r_i = (\psi(b'_1 \dots b'_i)^{-1}u, \varepsilon_C)$ for $1 \leq i \leq l$.

Note that $(p_0, \tau_1, p_1) \in \Delta_3$ as $\psi(b_1) \leq_p \psi(b_1 \dots b_j)u$, $\psi(b_1 \dots b_j)u = w = \varphi(a)$, and thus $\psi(b_1) \leq_p \varphi(a)$ with $\psi(b_1) \neq \varepsilon_C$. Further, for $1 < i \leq j$,

$$(p_{i-1}, t_i, p_i) = ((\psi(b_1 \dots b_{i-1})^{-1}w, \varepsilon_C), (\varepsilon_A, b_i), (\psi(b_1 \dots b_i)^{-1}w, \varepsilon_C)) \in \Delta_5,$$

as $\psi(b_1 \dots b_{i-1})^{-1}w \neq \varepsilon_C$, $\psi(b_i) \leq_p \psi(b_1 \dots b_{i-1})^{-1}w$, and

$$\psi(b_i)^{-1}\psi(b_1 \dots b_{i-1})^{-1}w = [\psi(b_1 \dots b_{i-1})\psi(b_i)]^{-1}w = \psi(b_1 \dots b_i)^{-1}w.$$

Similarly, $(r_0, \rho_i, r_1) = ((u, \varepsilon_C), (\varepsilon_A, b'_1), (\psi(b'_1)^{-1}u, \varepsilon_C)) \in \Delta_5$, as $u \neq \varepsilon_C$ and $\psi(b'_1) \leq_p u$. Moreover for $1 < i \leq l$,

$$(r_{i-1}, \rho_i, r_i) = ((\psi(b'_1 \dots b'_{i-1})^{-1}u, \varepsilon_C), (\varepsilon_A, b'_i), (\psi(b'_1 \dots b'_i)^{-1}u, \varepsilon_C)) \in \Delta_5,$$

as $\psi(b'_1 \dots b'_{i-1})^{-1}u \neq \varepsilon_C$, $\psi(b'_i) \leq_p \psi(b'_1 \dots b'_{i-1})^{-1}u$, and

$$\psi(b'_i)^{-1}\psi(b'_1 \dots b'_{i-1})^{-1}u = [\psi(b'_1 \dots b'_{i-1})\psi(b'_i)]^{-1}u = \psi(b'_1 \dots b'_i)^{-1}u.$$

Noting that $p_0 = \iota$, $p_j = q_0 = q_k = r_0$ and $r_l = (\varepsilon_C, \varepsilon_C)$, it follows that the concatenation of the sequence $(p_{i-1}, \tau_i, p_i)_{i=1}^{j}$, n copies of the sequence $(q_{i-1}, \sigma_i, q_i)_{i=1}^{k}$, and $(r_{i-1}, \rho_i, r_i)_{i=1}^{k}$ gives a path of length $j + kn + l$ in $\mathcal{A}_{\varphi,\psi}$ originating from ι and ending in $(\varepsilon_C, \varepsilon_C)$.

This path has label $\tau_1\tau_2 \dots \tau_j(\sigma_1\sigma_2 \dots \sigma_k)^n\rho_1\rho_2 \dots \rho_l$, and hence

$$\tau_1\tau_2 \dots \tau_j(\sigma_1\sigma_2 \dots \sigma_k)^n\rho_1\rho_2 \dots \rho_l \in \mathcal{L}(\mathcal{A}_{\varphi,\psi})$$

for all $n \in \mathbb{N}$. Thus by Theorem 5.3.1,

$$\pi(\tau_1\tau_2 \dots \tau_j(\sigma_1\sigma_2 \dots \sigma_k)^n\rho_1\rho_2 \dots \rho_l) \in \mathcal{I}$$

for all $n \in \mathbb{N}$. Hence by Lemma 5.1.1, it follows that $\Pi(\varphi, \psi)$ is not finitely generated, as it has infinitely many indecomposable elements.

For (ii) $\Rightarrow$ (i), suppose that $\mathcal{A}_{\varphi,\psi}$ has no cycles. As A and B are finite, it

follows that Q and δ are finite, and hence $\mathcal{A}_{\varphi,\psi}$ is a finite automaton. As $\mathcal{A}_{\varphi,\psi}$ has no cycles, then there are only finitely many paths in $\mathcal{A}_{\varphi,\psi}$, and thus the language accepted by $\mathcal{A}_{\varphi,\psi}$ is finite. By Theorem 5.3.1, $\Pi(\varphi,\psi)$ has finitely many decomposable elements, and hence is finitely generated by Lemma 5.1.1. □

For a given fiber product $\Pi(\varphi,\psi)$, the associated automaton $\mathcal{A}_{\varphi,\psi}$ can be viewed as a directed graph with finitely many nodes and edges. As it is decidable whether or not a finite directed graph has cycles, we have shown the following result.

Theorem 5.4.2. *Given a fiber product $\Pi(\varphi,\psi)$ of two free monoids A^* and B^* over the free monoid fiber C^* (where $|A|,|B|,|C| < \infty$), it is decidable whether or not $\Pi(\varphi,\psi)$ is finitely generated as a monoid.* □

Having dealt with the case for free monoids, we are able also derive analagous results for free semigroups. Given epimorphisms $\varphi : A^+ \to C^+$, $\psi : B^+ \to C^+$ (with A, B, C finite alphabets), we can extend φ and ψ naturally to epimorphisms $\varphi' : A^* \to C^*$, $\psi' : B^* \to C^*$ by mapping ε_A and ε_B to ε_C.

As $\Pi(\varphi',\psi') = \Pi(\varphi,\psi) \cup \{(\varepsilon_A,\varepsilon_B)\}$, then $\Pi(\varphi,\psi)$ is finitely generated as a semigroup if and only if $\Pi(\varphi',\psi')$ is finitely generated as a monoid. From Lemma 5.1.2, this gives us the following corollaries.

Corollary 5.4.3. *Let $\varphi : A^+ \to C^+$, $\psi : B^+ \to C^+$ be two epimorphisms with A, B, C finite alphabets, and let $\mathcal{A}_{\varphi',\psi'}$ be the associated automaton given in Definition 5.2.3, where $\varphi' : A^* \to C^*$ and $\psi' : A^* \to C^*$ are the natural monoid epimorphism extensions. Then the following are equivalent:*

(i) *$\Pi(\varphi,\psi)$ is finitely generated;*

(ii) *$\mathcal{A}_{\varphi',\psi'}$ is acyclic.* □

Corollary 5.4.4. *Given a fiber product $\Pi(\varphi,\psi)$ of two free semigroups A^+ and B^+ over the free semigroup fiber C^+ (where $|A|,|B|,|C| < \infty$), it is decidable whether or not $\Pi(\varphi,\psi)$ is finitely generated as a semigroup.* □

We conclude the chapter with some examples of finitely generated and non-finitely generated fiber products of free monoids over a free monoid.

Examples 5.4.5. (a) Recall from Examples 5.2.4 (a) that the fiber product

$$\Pi(\varphi, \psi) = \{(u, v) \in A^* \times B^* : |u| = |v|_a + 2|v|_b\}$$

has the associated two-tape automaton $\mathcal{A}_{\varphi,\psi}$ given by Figure 5.3, printed below.

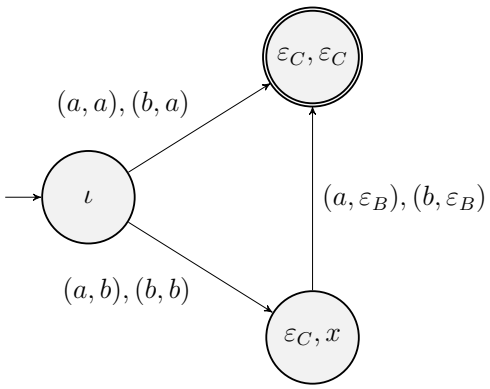


Figure 5.3: $\mathcal{A}_{\varphi,\psi}$ for $\Pi(\varphi, \psi)$ given in Examples 5.4.5 (a).

We saw in Example 4.3.8 that $\Pi(\varphi, \psi)$ was finitely generated. This is now evidenced once more by Theorem 5.4.1, as we can see by inspection of 5.3 that $\mathcal{A}_{\varphi,\psi}$ has no cycles.

(b) Recall from Examples 5.2.4 that the fiber product

$$\Pi(\varphi, \psi) = \{(u, v) \in A^+ \times B^+ : |u|_a + 2|u|_b = 2|v|_a + |v|_b\}$$

has the associated two-tape automata $\mathcal{A}_{\varphi,\psi}$ given by Figure 5.4, overleaf.

Then by Theorem 5.4.1, $\Pi(\varphi, \psi)$ is not finitely generated, as it has a cycle between the states (ε_C, x) and (x, ε_C). The automaton accepts words in the language

$$\{(a, a)[(b, \varepsilon_B)(\varepsilon_A, a)]^m(a, \varepsilon_B) \in \Sigma^* : m \in \mathbb{N}\}.$$

Considering the image of these words under the product monoid mapping π,

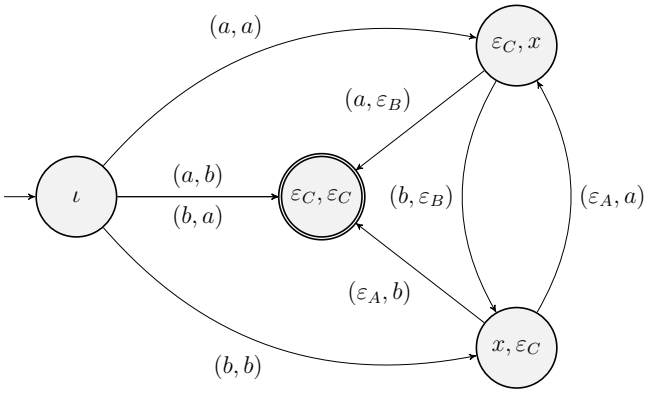


Figure 5.4: $\mathcal{A}_{\varphi,\psi}$ for $\Pi(\varphi,\psi)$ given in Examples 5.4.5 (b).

then by Theorem 5.3.1, it follows that the set

$$\{(ab^m a, a^{m+1}) \in A^* \times B^* : m \in M\}$$

is an infinite set of indecomposable elements of $\Pi(\varphi,\psi)$. $\triangle$

Chapter 6

Concluding remarks and further questions

In this concluding chapter, we briefly summarise the thesis by reflecting on our findings, and pose some arising questions for future research. A number of these questions have also been posed in [7] and [6], written by the author.

In Chapter 2, we commented that $\mathbb{Z}\times\mathbb{Z}$ had only two non-trivial subgroups up to isomorphism, and hence contained only finitely many subdirect products up to isomorphism. By contrast, we proved that the direct product $\mathbb{N} \times \mathbb{N}$ has uncountably many subsemigroups up to isomorphism (Theorem 2.1.12). This was later used to prove that the finite power $\mathbb{N}^k$ had uncountably many subdirect products up to isomorphism (Theorem 2.3.3).

Both $\mathbb{N}$ and $\mathbb{Z}$ are examples of finitely generated commutative semigroups, but have different subdirect substructure. We thus ask the following.

Question 6.0.1 ([7, Question 4.1])**.** Is it possible to characterise all pairs of finitely generated commutative semigroups S , T such that there are only countably many pairwise nonisomorphic subdirect products of S and T?

We also give a generalisation of this question to finitely generated free semigroups of other varieties, as below.

Question 6.0.2 ([7, Question 4.3])**.** How many pairwise non-isomorphic subsemigroups and subdirect products does $F \times F$ contain, where F is a finitely generated free semigroup in some other well-known semigroup varieties, such as inverse semigroups or completely regular semigroups?

We derived that the direct product of two semigroups, both of which with infinite order elements, has uncountably many subsemigroups up to isomorphism (Corollary 2.1.13). We ask the following additional questions based on this result.

Question 6.0.3. Does Corollary 2.1.13 generalise to subdirect products? That is, if S and T are semigroups with elements of infinite order, can we always find uncountably many subdirect products of $S \times T$ up to isomorphism? If not, for which S and T is this true?

We also classified in Theorem 2.2.4 and Theorem 2.4.1 the finite semigroups S for which $\mathbb{N} \times S$ has only countably many subsemigroups, and subdirect products respectively. We found that in both cases, the structure of S played the key role in determining countability. In the subsemigroups case, S needed to be a union of groups. In the subdirect product case, every element of S needed to have a relative left or right identity. We saw in Examples 2.4.2 that such semigroups can be as small as order 2. This is surprising, as we saw that $\mathbb{N}$ had only countably subsemigroups up to isomorphism, but $\mathbb{N} \times \{x, 0\}$ did not, for $\{x, 0\}$ the two element zero semigroup. In a similar vein to Question 6.0.1, we ask the folllowing.

Question 6.0.4 ([7, Question 4.2])**.** Given a finitely generated infinite commutative semigroup S, is it possible to characterise all finite semigroups T such that $S \times T$ has only countably many pairwise non-isomorphic subsemigroups or subdirect products? Do these characterisations depend on S?

Leading on from the results of Chapter 2, we continued our combinatorial survey of subdirect products of free semigroups in Chapter 3, with a view towards finite generation. In particular, we found the number of sets of letter pairs which generated subdirect products of free semigroups in Lemma 3.1.4. Within these sets, we then counted the number which also generated fiber

products in Corollary 3.2.5. The analytic comparison of Proposition 3.3.1 suggested that finitely generated subdirect products of this type were numerous, but we argued in Proposition 3.3.2 that few of them arise as fibered products.

Sets of letter pairs are just one example of finitely generating subdirect products however. Any finite generating set for a subdirect product of free semigroups has a bound on the length of the words in each factor. Sets of letter pairs are the special case where this bound is set to be one. We hence ask the following.

Question 6.0.5. Given a bound $k \in \mathbb{N}$, how many subsets of $A^+ \times B^+$ consisting of pairs (u, v) with $|u|, |v| \leq k$ generate a subdirect product of $A^+ \times B^+$? How many generate fiber products? Is it decidable when a given subset generates a fiber product? Is there a matrix characterisation of such subsets, as in the proofs of Lemma 3.1.4 and Corollary 3.2.5?

In Chapter 4, we concentrated on the finitary properties of finite generation and finite presentation, in particular for fiber products of free semigroups and monoids. We saw that for finite fibers, fiber products of free monoids are finitely generated only in the case where the fiber is a cyclic group, and the associated epimorphisms are constant on the alphabet (Theorem 4.1.4). In the case of finite generation however, we saw that we also have finite presentation, and gave such presentations in the proof of Theorem 4.2.1. We argued that this perhaps indicates that fiber products of free semigroups are hard to finitely generate.

Contrastingly, we showed that fiber products of free semigroups over finite fibers are never finitely generated or presented (Proposition 4.3.1). We saw that, for finitely generated fiber products of free semigroups, the fiber need be an infinite, $\mathcal{J}$-trivial, idempotent-free finitely generated semigroup (Proposition 4.3.1, Proposition 4.3.3, and Proposition 4.3.4). These conditions on a semigroup are particularly restrictive, but we saw in Lemma 4.3.6 that they are not sufficient to ensure finite generation. Additional requirements for the epimorphisms are also likely necessary. We ask the following questions based

on this material.

Question 6.0.6 ([6, Question 7.1])**.** Does there exist a finitely generated subdirect product (non-fibered) of two free monoids with a finite quotient, which is not finitely presented?

Question 6.0.7. Find further necessary conditions on a fiber product of free semigroups to be finitely generated. What are the necessary conditions for the Cayley graph of the fiber?

Question 6.0.8 ([6, Question 7.3])**.** Does Theorem 4.1.4 generalise to fiber products of free monoids over *infinite* fibers? That is, is every finitely generated fiber product of two free monoids also finitely presented?

We finished our study in Chapter 5, where we discussed the finite generation problem for fiber products of free semigroups and monoids, particularly with free fibers. We determined, in general, that the number of indecomposable elements of a fiber product ultimately decides whether it is finitely generated or not (Lemma 5.1.1, Lemma 5.1.2). We utilised a two-tape automatic construction in order to decide the cardinality of the set of indecomposables, showing that such automata accept a language in bijection with the indecomposables (Theorem 5.3.1). We deduced that this set was finite if and only if the automata constructed were acyclic (Theorem 5.4.1,Corollary 5.4.3), which is a decidable property of graphs. We concluded that the finite generation problem for fiber products of free semigroups and monoids over free fibers is decidable (Theorem 5.4.2, Corollary 5.4.4).

The following questions arise from these results.

Question 6.0.9. In the case of finite generation for a fiber product of two free semigroups/monoids over a free fiber, can we deduce a finite presentation for it, given the associated two-tape automaton?

Question 6.0.10. Can two-tape automata be used to recognise the indecomposable elements of fiber products over other infinite fibers? What properties of such automata can be deduced, and how do they relate to the structure of the fiber?

Bibliography

[1] I.M. Araújo, N. Ruškuc, "On finite presentability of direct products of semigroups". *Algebra Colloq.* **7** 83–91. (2000)

[2] G. Baumslag, J.E. Roseblade, "Subgroups of direct products of free groups". *J. London Math. Soc.* **30**, 44–52. (1984)

[3] M.R. Bridson, C.F. Miller III, "Structure and finiteness properties of subdirect products of groups". *Proc. Lond. Math. Soc.* **98**, 631—651. (2009)

[4] S. Burris, H.P. Sankappanavar, "A course in universal algebra". *Amer. Math. Monthly.* **78**, Springer-Verlag, New York. (1981)

[5] J.L. Chrislock, T. Tamura, "Notes on subdirect products of semigroups and rectangular bands". *Proc. Amer. Math. Soc.* **20**, 511–514. (1969)

[6] A. Clayton, "On finitary properties of fiber products for fiber products of free semigroups and free monoids". Semigroup forum, accepted. https://arxiv.org/abs/1907.01378. (2019)

[7] A. Clayton, N. Ruškuc, "On the number of subsemigroups of direct products involving the free monogenic semigroup". *J. Austral. Math. Soc.* First view, 12p. Available online. (2019)

[8] A.H. Clifford, G.B. Preston, "The Algebraic Theory of Semigroups, Vol. 1". *Surveys Amer. Math. Soc.* **7**. Providence, Rhode Island. (1961)

[9] A.H. Clifford, G.B. Preston, "The Algebraic Theory of Semigroups, Vol. 2". *Surveys Amer. Math. Soc.* **7**. Providence, Rhode Island. (1967)

[10] J.-A. de Séguier, "Éléments de la Théories de Groupes Abstraits". Gauthier-Villars, Paris. (1904)

[11] L.E. Dickson, "On semi-groups and the general isomorphism between infinite groups". *Trans. Amer. Math. Soc.* **6**, 205–208. (1905)

[12] R. Gray, N. Ruškuc, "On residual finiteness of direct products of algebraic systems". *Monatsh. Math.* **158**, 63–69. (2009)

[13] F. Grunewald, "On some groups which cannot be finitely generated". *J. London Math. Soc.* **17** (2), 427–436. (1978)

[14] M. Hall, Jr., "The Theory of Groups". *AMS.* Chelsea Publishing, 63–64. (1999)

[15] P.M. Higgins, "Techniques of Semigroup Theory". *Oxford University Press.* Oxford. (1992)

[16] J.M. Howie, "Fundamentals of Semigroup Theory". *LMS Monographs.* **12**, The Clarendon Press, New York. (1995)

[17] H.-K. Ju, S. Seo, "Enumeration of $(0,1)$-matrices avoiding some 2×2 matrices". *Discrete Math.* **312**, no. 16, 2473—2481. (2012)

[18] S. Lang, "Algebra". *Graduate texts in Mathematics.* **211**, 880–881 Springer-Verlag, New York (3rd edition). (2002)

[19] P. Mayr, N. Ruškuc, "Generating subdirect products". *J. London Math. Soc.* **100**, no. 2, 404–424. (2019)

[20] R. McKenzie, "Subdirect powers of non-abelian groups". *Houston J. Math.* **8**, no. 3, 389–399. (1982)

[21] K.A. Mikhaĭlova, "The occurence problem for direct products of groups". *Mat. Sb. (N.S.)* **70(112)**, 241–251. (1966)

[22] H. Mitsch, "Subdirect products of E-inverse semigroups". *J. Austral. Math. Soc.* **48**, 66–78. (1990)

[23] K.S.S. Nambooripad, R. Veeramony, "Subdirect products of regular semigroups". *Semigroup Forum.* **28**, 265–307. (1983)

[24] J.E. Pin, "Varieties of formal lanuages". *North Oxford Academic Press.* North Oxford, London (English translation). (1986)

[25] D. Rees, "On semi-groups". *Proc. Cam. Math. Soc.* **3**, 387–400. (1940)

[26] E.F. Robertson, N. Ruškuc, J. Wiegold, "Generators and relations of direct products of semigroups". *Trans. Amer. Math. Soc.* **350**, 2665–2685. (1998)

[27] W. Sit, M-K. Siu, "On the subsemigroups of $\mathbb{N}$". *Math. Mag.* **48**, 225—227. (1975)

[28] A.K. Sushkevich, "Über die endlichen Gruppen ohne das Gesetz der eindeutigen Umkehrbarkeit". *Math. Ann.* **99**, 30–50. (1928)

[29] V.V. Wagner, "Generalised groups". *Proc. USSR Acad. Sci* (Russian). **84**, 1119–1122. (1952)

Acknowledgements

I have many, many people to thank for their input during my time at St Andrews. I hope to not have forgotten anyone in the twilight of these pages, especially those who have made it this far. I would like to thus start by covering all possibilities, and thank absolutely everybody I have met over the past three and a half years.

In terms of specific thanks, I would like to start by giving a sincere, heartfelt and full thank you to my primary supervisor, Professor Nik Ruškuc. This work simply could not have been undertaken without his guidance; academically, personally, humorously or otherwise. That I have come through to the other end is great testament to his skill and patience as a supervisor. Thank you for always being incredibly helpful, honest, and above all, a great pleasure to work with. I hope to represent our research publically with fast food analogies for many more years to come.

I would also like to thank my second supervisor, Professor James Mitchell. James has been a great source of help, especially in terms of funding. He was also one of the first people to meet with me when I first visited St Andrews, and his great humour and personable manner very much sold me on the University. I apologise for any hair related (or rather, no hair related) jokes I may have made over the years.

I would like to thank the EPSRC (grant number EP/N509759/1) for their generous funding and skills training. I would also like to thank the entire School of Mathematics & Statistics of the University of St Andrews for their financial support, training, teaching opportunities, administrative aid, and for being genuinely an extraordinarily friendly place in which to have spent the past three and a half years. Particular thanks goes to the secretarial staff, who have helped me administratively during the entire process. Many thanks to Valerie Sturrock, Niki

Stalker, and Lauren Gatherum for kindly halting their busy schedules now and then to engage with my gossipy procrastination.

Without the emotional support of the friends I have made here at St Andrews, I would simply not have had the mental ability to complete this work. I would particularly like to thank my best friend, Gerry O'Reilly, for making my time in St Andrews more than memorable. Alongside being a fiercely intelligent friend with whom I can discuss the research we both interest ourselves in, Gerry is one of the funniest people I know and always manages to make me smile in times when I would perhaps not otherwise. Our many social engagements have given me an incredible number of memories, and I thank him deeply for his time and close friendship. I hope to be ever amazed by his French Horn playing for many years to come.

I would also like to thank Ben Williams, who had been a tremendous amount of help in getting to know St Andrews when he was a student here. Ben has been a particular help in the completion of this thesis, having gone through the mill prior to me, and urging me to keep going with his "deadline countdowns". He has more importantly been a great friend, and I thank him very much for the time spent socially when he was here.

I'd further like to thank, in no particular order, my good friends from "down the corridor" in the school; Veronica Kelsey, Chiara Villa, Raad Al Kohli, Matt McDevitt, Finn Smith, Fanny Empacher, Nayab Khalid, Douglas Howroyd, Lawrence Lee, Craig Miller, Fernando Flores Brito, Tom Howson, Craig Johnston, and the entire Solar PhD group. Though this a long corridor stretching multiple floors (and buildings), all of these people amongst others have visited me and spent may hours together in my office, making the entire PhD process a friendly and engaging one. Moreover, many of these friendships have extended much much further than the corridor, and I am deeply grateful to all of them for the times we have spent together outside of the office, as much as I am for the time spent in.

Finally, I would like to thank my family, who have supported me in numerous ways during this thesis. To my mum Linda, dad Charles, sister Alex, cousins Paul, Dawn & Sophie, and Uncle Bubby; thank you for your patience and support during my time away in Scotland. Also to my loving late grandparents, Alan & Nora, for their support in the early years of my education.